Australian Government
Department of Defence
Defence Science and
Technology Organisation

Comparative Descriptive Statistics of Skewed Probability Distributions

M.P. Fewell

Maritime Operations Division
Systems Sciences Laboratory

DSTO-TR-1596

ABSTRACT

This report is a brief handbook on the comparative descriptive statistics of a wide variety of skewed probability distributions, both continuous and discrete. The aim is to facilitate the comparison of different distributions, for use where random variables are employed without any firm information on their distribution. In this situation, it is of interest to look for sensitivity to the distribution chosen. This can best be done by running the model with a variety of distributions, which then raises the question of how to compare distributions. This work advocates the use of moments and presents the requisite equations. As obvious as this approach may appear, many of the equations do not seem to have been published previously and some of the results are apparently wholly new. A total of 18 distributions are treated in detail, including all of the most commonly used skewed probability distributions.

RELEASE LIMITATION

Approved for public release

Published by

DSTO System Sciences Laboratory
PO Box 1500
Edinburgh South Australia 5111 Australia

Telephone: (08) 8259 5555
Fax: (08) 8259 6567

AR-013-136
July 2004

APPROVED FOR PUBLIC RELEASE

Australian Government
Department of Defence
Defence Science and
Technology Organisation

DSTO

Comparative Descriptive Statistics of Skewed Probability Distributions

M.P. Fewell

DSTO-TR-1596

Comparative Descriptive Statistics of Skewed Probability Distributions

Executive Summary

This Report addresses a technical issue in the current practice of operational analysis (OA). Increasingly, OA studies involve simulations of varying levels of sophistication. A feature of all simulations is the use of random variables, and this immediately raises the question of what distribution to employ. Should the random variable be taken as uniformly distributed over some range, or should it follow a bell curve with some mean and standard deviation, or be exponentially distributed, and so on? Sometimes there are methodological or theoretical arguments favouring a particular distribution, but often there are not. In their absence, the modeller is forced to make a more-or-less arbitrary choice of distribution, motivated perhaps by ease of use or personal familiarity. It then becomes interesting to know how sensitive the results of the analysis are to the choice of distribution: if another had been chosen, would the conclusions have been different?

The obvious way of approaching this question is to run the model using a variety of distributions. One must not, however, substitute one distribution for another uncritically. For example, usually the results of a model will be affected by the mean of an input variable, so one must match means when changing distributions. However, most common distributions have two parameters, so the requirement of matching means is not in itself sufficient. This Report advocates matching distribution moments: for two-parameter distributions, one should match means and standard deviations; with three parameters, the skewness should also be matched, and so on. As obvious as this approach may seem, the requisite equations have not been encountered in any of the texts on probability distributions examined for this work.

Hence, this Report gathers together equations for distribution parameters corresponding to a given mean, standard deviation and, where required, higher moments of a wide range of skewed probability distributions. Most of the equations were derived during the course of this work; many appear not to have been published before. All of the common skewed distributions are treated in detail, including

- 3 one-parameter distributions
- 11 two-parameter distributions
- 2 three-parameter distributions
- 1 four-parameter distribution
- 9 continuous distributions with the range zero to infinity
- 3 continuous distributions with a finite range (a to b)
- 3 discrete distributions with a finite range (a to b).

In addition, 11 lesser-known discrete distributions with finite ranges are mentioned, to give a flavour of the diversity available.

The emphasis on finite-range and semi-infinite ($0 \rightarrow \infty$) skewed distributions arose from the original motivation for this work, which was to support the modelling of decision times and times to carry out a given task. There is no obvious interpretation for a negative decision time, so it is natural to use distributions that are positive definite when modelling such quantities. However, the Gaussian distribution (defined from $-\infty$ to ∞) is also included because it is so widely known and used.

Author

M.P. Fewell
Maritime Operations Division

Matthew Fewell joined DSTO in 2001, coming from an academic physics background. He has worked and published in the fields of experimental nuclear structure physics, gaseous electronics, atom–photon interactions including coherent effects, laser physics and the plasma processing of materials. His present interests include military experimentation, particularly human-in-the-loop experiments in the maritime domain, and modelling the effects of network centricity. This has led him into the area of command and control and made him an interested spectator of research into decision making.

Contents

List of Figures

List of Tables

1. Introduction

In many, perhaps most, studies using random variates, the question of which distribution to use is not a significant issue: theoretical or methodological arguments point to the right one for the purpose. The issue of comparison between distributions does not then arise. In some cases, however, there is no clearly preferred probability distribution; the choice is more or less arbitrary and this raises the question of the sensitivity of the analysis to that choice. In these circumstances it is important to have a rational basis for comparing distributions, so that one is comparing like with like as far as practicable. The method advocated in this report is to match moments of distributions: where distributions have a single parameter, this is chosen to equate means; for two-parameter distributions, means and standard deviations are matched, and so on. This seems an entirely obvious approach, yet it is not mentioned in any of the many references examined in the course of this work. As a consequence, the requisite equations are not compiled anywhere, so this work was carried out to fill the gap. Some of the results presented herein are available in texts; most are not. For example, nothing like Figures 2.5 or 3.4–3.6 has been encountered in the literature.

Most of the distributions considered in this report have two parameters. Equations were derived and are presented herein for parameter values that give a desired mean and standard deviation. The differences between distributions are then quantified by skewness η_3 and kurtosis η_4, which are respectively the third and fourth moments about the mean divided by the third and fourth powers of the standard deviation. Expressions for skewness and kurtosis in terms of mean and standard deviation are presented where possible. Often, the expressions are more or less straightforward re-arrangements of standard results, but in some cases the equations have no closed-form solution; for these, graphs and tables are presented that assist in the choice of parameters.

In addition to the two-parameter distributions, three one-parameter, two three-parameter and one four-parameter distributions are included. For these, equations are presented giving the parameters in terms of as many moments as required.

The focus on skewed distributions arose from the original motivation for this work, which was to support a study on modelling decision making [1]. The time taken to make a decision has a natural definition for positive values only; hence, distributions of decision time ought to be strictly zero for $t < 0$. Further, it has been argued that the likelihood of an instantaneous decision is negligible, so the distribution of decision times ought to be zero at $t = 0$ as well. Most distributions treated herein have these characteristics; the two exceptions (Gaussian and exponential) are included because they are so widely used. Recently, the results reported herein have been applied to a study of service times in a queueing-theory model of maritime interception [2]. This is another case where one requires the distribution to be zero for $t \leq 0$.

This work was prompted by studies from the Naval Undersea Warfare Center Division Newport advocating the use of the inverse Gaussian probability density as a representation for decision times [3,4]. This distribution is not widely known and is omitted from several well known texts (e.g. [5–8]) although it was first described over half a century ago [9] and its properties have been documented for several decades [10 (ch.15),11]. In §2 of this report, its properties are compared with those of many other probability-density distributions. The emphasis naturally falls on other distributions on $[0,\infty)$, although the Gaussian distribution is also included because it is so ubiquitous and to provide a point of comparison with a non-skewed distribution.

The property that makes the inverse Gaussian distribution particularly suited to the description of decision speed, the behaviour of the 'decision rate' or 'hazard rate' at long times, is highlighted in §2.3.4.

As well as speed, decisions are characterised by soundness [1]. Although perhaps not essential, it is common for soundness scales to be finite: 0–1 say, or 0–10. The modelling of soundness is just as likely to be probabilistic as that of decision speed, so finite-range distributions are described in §3. Often, soundness is ranked on a discrete scale, 0–3, 1–5 etc., the so-called 'Likert scales'. This is also amenable to a probabilistic treatment; appropriate probability distributions are described in §3.3.

This work serves as a brief handbook on the comparative descriptive statistics of a wide variety of skewed probability distributions, both continuous and discrete. The aim is to facilitate the comparison of different distributions.

2. Semi-Infinite Distributions

As mentioned in §1, the inverse Gaussian probability density has been advocated for representing decision speed. The basic properties expected of a distribution of decision speeds are as follows:

- The domain of the distribution must not include negative times t.
- The probability density should be zero at $t = 0$; for otherwise there is a non-zero probability of an instantaneous decision. In particular, the exponential distribution, which has maximum probability density at $t = 0$, is not suitable.
- Whether the domain of the probability density should extend to infinity is not clear, but there are as yet no data establishing an upper limit on decision time. Thus, the domain of the distribution should be $0 \le t \le \infty$, so as to avoid introducing an arbitrary parameter.

The inverse Gaussian probability distribution, among others, fulfils these criteria. This section presents the main properties of this distribution and compares it with other distributions also satisfying the properties listed above. The feature distinguishing the inverse Gaussian probability density from the rest—the behaviour of its decision rate—is addressed in §2.3.4.

2.1 The Inverse Gaussian Distribution

In standard form, the definition of the inverse Gaussian probability density is [10–12]

$$f_{\mathrm{IG}}(t) = \sqrt{\frac{\lambda}{2\pi t^3}} \exp\left[-\frac{\lambda(t-\mu)^2}{2\mu^2 t}\right], \tag{2.1}$$

where the parameters λ, μ must both be greater than zero. As noted above, the domain is $0 \le t \le \infty$. By direct calculation, it is found that the mean of the distribution is μ and its variance σ^2 equals μ^3/λ. Since the function has two parameters, we re-write it using the mean μ and standard deviation σ as parameters. This is straightforward and has the added advantage that a little rearrangement simplifies the argument of the exponential function:

$$f_{\mathrm{iG}}(t) = \sqrt{\frac{\mu}{2\pi t^3}}\, \mathrm{v} e^{\mathrm{v}^2} \exp\left[-\frac{\mathrm{v}^2}{2}\left(\frac{t}{\mu} + \frac{\mu}{t}\right)\right], \tag{2.2}$$

where $\mathrm{v} = \mu/\sigma$.

Figure 2.1 shows plots of the distribution for three values of σ/μ. As Equation (2.2) indicates, the argument of the distribution scales with μ, so it is useful to plot Figure 2.1 with t/μ on the abscissa. Scaling is made complete by plotting the product $\mu f_{\mathrm{iG}}(t)$ and labelling the curves with the 'coefficient of variation' $s = \sigma/\mu = 1/\mathrm{v}$.

As may be expected, and as Figure 2.1 shows, the inverse Gaussian distribution becomes more symmetric as $\sigma/\mu \to 0$ and more skewed as $\sigma/\mu \to \infty$. The mode t_{m} (formula in Table 2.1 below) lies between zero and the mean μ for all values of σ/μ, with $t_{\mathrm{m}} \to 0$ as $\sigma/\mu \to \infty$. The probability density $f_{\mathrm{iG}}(t_{\mathrm{m}})$ at the mode rises as either $\sigma/\mu \to 0$ or $\sigma/\mu \to \infty$. The minimum value of $\mu f_{\mathrm{iG}}(t_{\mathrm{m}})$ occurs when $\sigma/\mu = 1/\sqrt{2} \approx 0.7071$. (This does not depend upon μ.)

As mentioned in §1, the method of comparing the various distributions is to examine functions with the same values of μ and σ; differences are then be expressed in terms of the higher-moment ratios, skewness η_3 and kurtosis η_4. Of course, yet higher moments exist, but η_3 and η_4 are usually sufficient to display trends. For the inverse Gaussian distribution, direct evaluation of the integrals concerned gives simple expressions for $\eta_{3\mathrm{iG}}$ and $\eta_{4\mathrm{iG}}$ in terms of μ and σ:

$$\begin{aligned} \eta_{3\mathrm{iG}} &= 3\sqrt{\mu/\lambda} = 3\sigma/\mu \\ \eta_{4\mathrm{iG}} &= 3 + 15\mu/\lambda = 3 + 15(\sigma/\mu)^2. \end{aligned} \tag{2.3}$$

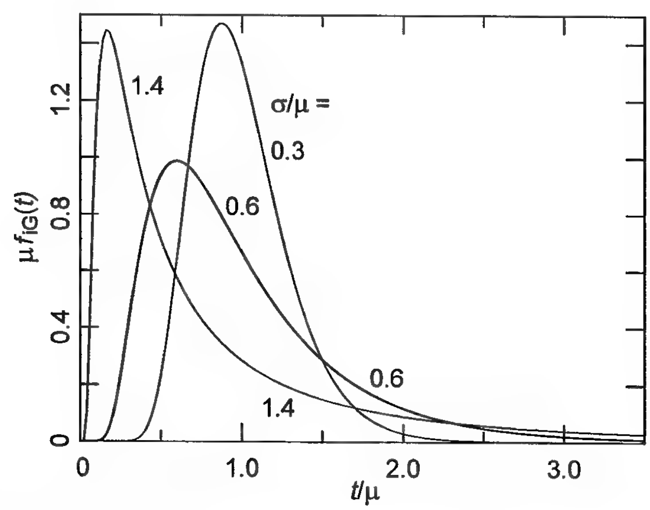

Figure 2.1: Examples of the inverse Gaussian distribution. Curves are labelled by values of the coefficient of variation σ/μ.

2.2 Other Distributions

This section discusses the properties of the Gaussian distribution and eight semi-infinite distributions: exponential, log-normal, gamma, bi-exponential, Weibull, log-logistic, inverted gamma and beta-prime. The list comprises all the best known distributions on $[0,\infty)$, including all those listed in a widely used text on simulation [7(pp. 299–318)]. The Gaussian and exponential distributions are not realistic candidates for describing decision speed, but are included because they are so common. The Gaussian also provides a point of reference to a non-skewed distribution.

In the following Section, the distributions are compared using the mean μ and standard deviation σ as the parameters, as in Equation (2.2). Hence, the aim of this Section is to invert the expressions for μ and σ in terms of the standard parameters. This can be done in closed form in all but two cases (Weibull and log-logistic). For these two, graphs and tables are presented to facilitate the transformation. Details of the inversion are presented in the following subsections. For convenience, expressions for the main properties of the distributions are collected in Table 2.1 (pp. 6–7).

2.2.1 Gaussian Distribution

The standard form of the Gaussian distribution already uses the mean μ and standard deviation σ as parameters:

$$f_G(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-(t-\mu)^2}{2\sigma^2}\right]. \tag{2.4}$$

2.2.2 Exponential Distribution

This distribution has just one parameter, the mean μ, which must be positive. The standard form is

$$f_e(t) = \frac{1}{\mu} \exp\left(\frac{-t}{\mu}\right) \qquad (0 \le t < \infty). \tag{2.5}$$

The standard deviation of this distribution equals μ also.

2.2.3 Log-Normal Distribution

The standard form of the log-normal distribution is often written using the symbols μ and σ to represent the mean and standard deviation of a notional underlying Gaussian distribution (e.g. [6,7,12,13]), rather than of the log-normal distribution itself. This confusing notation can be simply avoided by adopting other symbols. For example, Olkin *et al.* [5] write

$$f_{\text{l-n}}(t) = \frac{1}{t\sqrt{2\pi\delta^2}} \exp\left[\frac{-(\ln t - \xi)^2}{2\delta^2}\right] \qquad (0 \le t < \infty) \tag{2.6}$$

where δ, ξ are the parameters, both of which must be positive. The mean μ and standard deviation σ *of the log-normal distribution* are given by

$$\mu = \exp\left(\xi + \delta^2/2\right), \qquad \sigma = \exp\left(\xi + \delta^2/2\right)\sqrt{\exp(\delta^2) - 1}\,. \tag{2.7}$$

These equations can be inverted, leading to the following expression for the distribution in terms of μ, σ:

$$f_{\text{l-n}}(t) = \frac{1}{t\sqrt{2\pi \ln(s^2+1)}} \exp\left[-\frac{\ln^2\left(t\sqrt{s^2+1}/\mu\right)}{2\ln(s^2+1)}\right], \tag{2.8}$$

where $s = \sigma/\mu$, the inverse of ν.

2.2.4 Gamma Distribution

In standard form, the gamma distribution reads [5]

$$f_\Gamma(t) = \frac{\theta^r}{\Gamma(r)} t^{r-1} e^{-\theta t} \qquad (0 \le t < \infty), \tag{2.9}$$

where the parameters r, θ must both be positive. $\Gamma(x)$ is the usual gamma function [8 (§6.1.1)]. The equations for the mean μ and standard deviation σ in terms of the parameters,

$$\mu = r/\theta, \qquad \sigma^2 = r/\theta^2, \tag{2.10}$$

are easily inverted, so the distribution can be rewritten using these as parameters:

$$f_\Gamma(t) = \frac{\nu^{\nu^2}}{\sigma^{\nu^2}\Gamma(\nu^2)} t^{\nu^2-1} \exp\left(-\frac{\nu t}{\sigma}\right) \tag{2.11}$$

where $\nu = \mu/\sigma$, as for the inverse Gaussian distribution.

To obtain $f_\Gamma(0) = 0$, as required, one must take $r > 1$. This implies $\mu/\sigma > 1$. That is, it is not possible to have the standard deviation larger than the mean and still retain $f_\Gamma(0) = 0$. This is in contrast to the behaviour of the inverse Gaussian and log-normal distributions.

2.2.5 Bi-exponential Distribution

As noted above, the exponential distribution is not suitable for representing decision speed because its maximum value lies at $t = 0$. This defect can be rectified by a simple generalisation:

$$f_{\text{b-e}}(t) = \frac{\alpha\beta}{\beta - \alpha}\left(e^{-\alpha t} - e^{-\beta t}\right) \qquad (0 \le t < \infty), \tag{2.12}$$

where the two parameters α, β must be non-equal and both positive, and one can take $\beta > \alpha$ without loss of generality. This is known as the bi-exponential distribution, also as the two-stage hypo-exponential or generalised Erlang distribution [14(p.448)]. Although this distribution has a simple form, expressions for the mean and standard deviation in terms of the parameters are rather more complicated than for the distributions treated so far:

$$\mu = \frac{\alpha+\beta}{\alpha\beta}, \qquad \sigma = \frac{\sqrt{\alpha^2+\beta^2}}{\alpha\beta}. \tag{2.13}$$

Equations (2.13) can be inverted:

$$\alpha = \frac{2/\mu}{1+\sqrt{2s^2-1}}, \qquad \beta = \frac{2/\mu}{1-\sqrt{2s^2-1}} \tag{2.14}$$

($s = \sigma/\mu$, as before), but is not expedient to substitute these equations into Equation (2.12). However, relatively simple expressions can be obtained for the mode, skewness and kurtosis in terms of s, as shown in Table 2.1.

*Table 2.1: Properties of several probability density functions in terms of their mean μ and standard deviation σ.**

	Inverse Gaussian	Log-normal	Gamma	Weibull	Bi-exponential	Gaussian	Exponential
Mode t_m	$\mu\left(\sqrt{1+\frac{9s^4}{4}}-\frac{3s^2}{2}\right)$	$\frac{\mu}{(s^2+1)^{3/2}}$	$\mu(1-s^2)$	$\alpha\sqrt[\beta]{\frac{\beta-1}{\beta}}$	$\frac{\mu}{2}\frac{1-s^2}{\sqrt{2s^2-1}}\ln\left(\frac{s^2+\sqrt{2s^2-1}}{1-s^2}\right)$	μ	0
Skewness η_3	$3s$	$3s+s^3$	$2s$	Eqn (2.18)	$\frac{3}{s}-\frac{1}{s^3}$	0	2
Kurtosis η_4	$3+15s^2$	$3+16s^2+15s^4$ $+6s^6+s^8$	$3+6s^2$	Eqn (2.19)	$6+\frac{6}{s^2}-\frac{3}{s^4}$	3	9
Behaviour at $t=0$	$f(0)$ and all its derivatives are zero	$f(0)$ and all its derivatives are zero	$f(0)$ and its first $\lfloor v^2-1\rfloor$ derivatives are zero†	$f(0)$ and its first $\lfloor\alpha-1\rfloor$ derivatives are zero†	$f(0)=0$; all derivatives are non-zero	$f(0)$ is non-zero	$f(0)$ is non-zero
Decision rate‡ $h(t)$	Eqn (2.46)	$\frac{f_{\text{l-n}}(t)}{1-\Phi[(\ln t-\xi)/\delta]}$ §	$\frac{f_\Gamma(t)\,\Gamma(v^2)}{\Gamma(v^2,tv^2/\mu)}$ #	$\frac{\beta t^{\beta-1}}{\alpha^\beta}$	$\alpha\beta\frac{e^{-\alpha t}-e^{-\beta t}}{\beta e^{-\alpha t}-\alpha e^{-\beta t}}$	$\frac{f_G(t)}{1-F_G(t)}$ ¶	$1/\mu$
$h(\infty)$‡	$\mu/(2\sigma^2)$	0	μ/σ^2	∞ (for $\beta>1$)	α	∞	$1/\mu$

* Coefficient of variation $s=\sigma/\mu=1/v$; closed-form expressions in terms of μ, σ not obtainable for Weibull or log-logistic distributions (see text).

† $\lfloor x\rfloor$ = largest integer $\le x$.

‡ See §2.3.4.

§ $\Phi(x)$ is the cumulative probability function¶ of a Gaussian distribution with zero mean and unit standard deviation.

¶ See Equation (2.45).

\# $\Gamma(a,x)$ is the complementary incomplete gamma function [8(§6.5.3)].

‖ $\gamma(a,x)$ is the incomplete gamma function [8(§6.5.2)].

▲ Use Equations (2.35) to determine the parameter values.

▼ $B(q,p,x)$ is the incomplete beta function, called $B_x(q,p)$ in Ref. 8 (§6.6.1).

Table 2.1 continued (see previous page for footnotes).

	Log-logistic	Inverted gamma	Beta prime▲
Mode t_m	$\beta\sqrt[\alpha]{\dfrac{\alpha-1}{\alpha+1}}$	$\mu\dfrac{s^2+1}{3s^2+1}$	$\tau\dfrac{p-1}{q+1}$
Skewness η_3	Eqn (2.24)	$\dfrac{4s}{1-s^2}$	(a parameter)
Kurtosis η_4	Eqn (2.25)	$\dfrac{3(1+7s^2)}{(1-s^2)(1-2s^2)}$	Eqn (2.37)
Behaviour at $t=0$	$f(0)$ and its first $\lfloor\alpha-1\rfloor$ derivatives are zero†	$f(0)$ and all its derivatives are zero	$f(0)$ and its first $\lfloor p-1\rfloor$ derivatives are zero†
Decision rate‡ $h(t)$	$\dfrac{\alpha\beta^{2\alpha}t^{\alpha-1}}{(t^\alpha+\beta^\alpha)^3}$	$\dfrac{[\mu(v^2+1)]^{v^2+2}e^{-\mu(v^2+1)/t}}{\gamma\left(v^2+2,\mu\dfrac{v^2+1}{t}\right)}$ ‖	$\dfrac{\tau^q t^{p-1}}{(t+\tau)^{p+q}B\left(q,p,\dfrac{\tau}{t+\tau}\right)}$ ▼
$h(\infty)$‡	0	0	∞

Equations (2.14) imply that $1/\sqrt{2} < \sigma/\mu < 1$, so that α and β are real and positive. Hence, as with the gamma distribution, the range of standard deviations that can be obtained for a given mean is restricted; the range available with the bi-exponential distribution is particularly small, as illustrated in Figure 2.5 below (p. 14).

2.2.6 Weibull Distribution

The standard form of the Weibull distribution is [5]

$$f_W(t)=\frac{\beta}{\alpha^\beta}t^{\beta-1}e^{-(t/\alpha)^\beta} \qquad (0\le t<\infty) \tag{2.15}$$

where the parameters α, β are both positive and $\beta>1$ is required to obtain $f_W(0)=0$.(a) As with other distributions, expressions for the mean and standard deviation in terms of the parameters can be obtained, but in this case they cannot be inverted in closed form. The expressions are:

$$\mu=\alpha\Gamma\left(1+\frac{1}{\beta}\right), \qquad \sigma=\alpha\sqrt{\Gamma\left(1+\frac{2}{\beta}\right)-\Gamma^2\left(1+\frac{1}{\beta}\right)}. \tag{2.16}$$

Since we consider only $\beta>1$ and since $\Gamma(x)<1$ in the range $1<x<2$, with a minimum at $\Gamma(1.46163)=0.885603$ [8(p.259)], it follows that $0.8856\alpha<\mu<\alpha$. That is, the mean lies within 12% of the value of α.

It is clear from Equations (2.16) that the ratio $s=\sigma/\mu$ depends on β alone. Using the recurrence and duplication formulae for the gamma function [8(p.256)], one obtains

$$s^2=\left(\frac{\sigma}{\mu}\right)^2=\frac{2^{2/\beta}\Gamma\left(\frac{1}{2}+\frac{1}{\beta}\right)}{\sqrt{\pi}\,\Gamma\left(1+\frac{1}{\beta}\right)}-1. \tag{2.17}$$

(a) Note that α and β are often interchanged, e.g. [6,7].

From this and the properties of $\Gamma(x)$, one sees that $s \to 0$ as $\beta \to \infty$ and $s \to 1$ as $\beta \to 1$. That is, as with the gamma and bi-exponential distributions, it is not possible to obtain $\sigma > \mu$ while retaining the desired behaviour near $t = 0$.

Expressions similar to Equations (2.16) for skewness η_3 and kurtosis η_4 read:

$$\eta_3 = \frac{\Gamma\left(1+\frac{3}{\beta}\right) - 3\Gamma\left(1+\frac{2}{\beta}\right)\Gamma\left(1+\frac{1}{\beta}\right) + 2\Gamma^3\left(1+\frac{1}{\beta}\right)}{\left[\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma^2\left(1+\frac{1}{\beta}\right)\right]^{3/2}}, \tag{2.18}$$

$$\eta_4 = \frac{\Gamma\left(1+\frac{4}{\beta}\right) - 4\Gamma\left(1+\frac{3}{\beta}\right)\Gamma\left(1+\frac{1}{\beta}\right) + 6\Gamma\left(1+\frac{2}{\beta}\right)\Gamma^2\left(1+\frac{1}{\beta}\right) - 3\Gamma^4\left(1+\frac{1}{\beta}\right)}{\left[\Gamma\left(1+\frac{2}{\beta}\right) - \Gamma^2\left(1+\frac{1}{\beta}\right)\right]^2}. \tag{2.19}$$

There seems little useful that can be done to reduce these equations further. It can be seen from these equations and those in Table 2.1 that α acts purely as a scaling parameter.

Figure 2.2 shows plots of properties of the Weibull distribution as a function of the shape parameter β. The practical value of this figure lies in its facilitation of the inversion of Equations (2.16) to obtain β for given values of σ/μ, a necessary step in the comparison process. Figure 2.2 shows that $1/\beta \approx \sigma/\mu$, but that σ/μ slightly exceeds $1/\beta$ except at zero and unity. Table 2.2 lists values of $1/\beta$ to six decimal places, obtained by numerical inversion of Equation (2.17), for selected values of σ/μ.

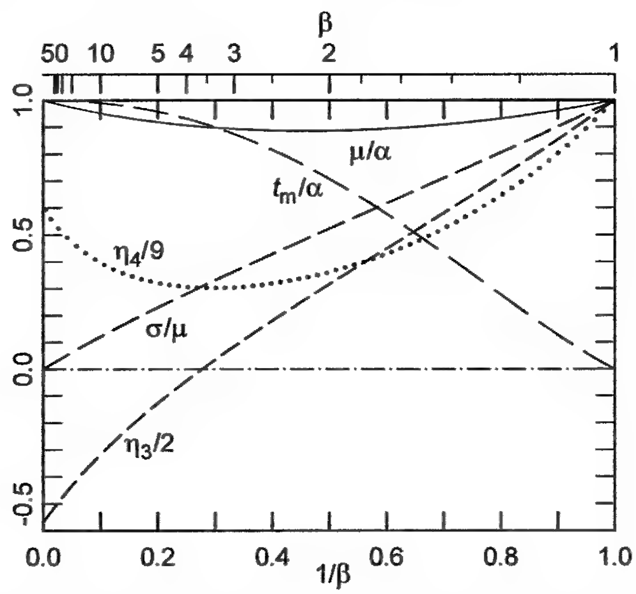

Figure 2.2: Various properties of the Weibull distribution as a function of the parameter β for the case $\beta > 1$. The curves show the ratios t_m/α, μ/α and σ/μ, where t_m is the mode, μ the mean and σ the standard deviation, and values of the skewness η_3 and kurtosis η_4. The skewness and kurtosis have been scaled by their respective values at $\alpha = 1$ to give convenient ranges of values for plotting.

Table 2.2: Values of 1/β to 6 decimal places for various values of σ/μ.

σ/μ	$1/\beta$	σ/μ	$1/\beta$	σ/μ	$1/\beta$	σ/μ	$1/\beta$
0.05	0.040081	0.30	0.269268	0.50	0.475885	0.75	0.742087
0.10	0.082281	1/3	0.302707	0.55	0.529030	0.80	0.794755
0.15	0.126471	0.35	0.319612	0.60	0.582383	0.85	0.846958
0.20	0.172491	0.40	0.370972	0.65	0.635775	0.90	0.898612
0.25	0.220157	0.45	0.423133	0.70	0.689055	0.95	0.949645

2.2.7 Log-Logistic Distribution

The logistic distribution has a $\mathrm{sech}^2\, t$ form, which has obvious similarities with the Gaussian distribution. The analogy suggests the definition of the log-logistic distribution: it has the same relationship to the logistic distribution as the log-normal has to the Gaussian [10(§23.11)]. The standard form of the density function is [7]

$$f_{l\text{-}l}(t) = \alpha\beta^{\alpha}\frac{t^{\alpha-1}}{\left(t^{\alpha}+\beta^{\alpha}\right)^2} \qquad (0 \le t < \infty), \tag{2.20}$$

where, for the general distribution, the parameters α, β are both positive. However, as with several other distributions, the requirement that $f_{l\text{-}l}(0) = 0$ imposes an additional restriction on one of the parameters. In this case, one must have $\alpha > 1$.

Yet further restrictions on α are placed by requirements that the moments of the distribution be finite. The rth moment μ_r about the origin is given by

$$\mu_r = \beta^r \int_0^{\infty} \frac{y^{r/\alpha}}{(1+y)^2}\,\mathrm{d}y\,, \tag{2.21}$$

an integral that is infinite unless $\alpha > r$. Hence, a finite standard deviation requires $\alpha > 2$, a finite skewness $\alpha > 3$, and so on. Since we plan to compare log-logistic with other distributions by specifying values of mean and standard deviation, it follows that we are interested only in those distributions with $\alpha > 2$.

Equation (2.21) is a standard integral [15(¶856.07)]; its evaluation for $r = 1, 2$ leads to the mean and variance:

$$\mu = \frac{\pi\beta}{\alpha}\mathrm{cosec}\frac{\pi}{\alpha}, \qquad \sigma^2 = \frac{\pi\beta^2}{\alpha}\left[2\,\mathrm{cosec}\frac{2\pi}{\alpha} - \frac{\pi}{\alpha}\mathrm{cosec}^2\frac{\pi}{\alpha}\right]. \tag{2.22}$$

As with the Weibull distribution, these equations cannot be inverted in closed form. The value of α is the solution to the equation

$$s = \frac{\sigma}{\mu} = \sqrt{\frac{\alpha}{\pi}\tan\frac{\pi}{\alpha} - 1}\,; \tag{2.23}$$

this is shown in Figure 2.3 as the curve labelled 'σ/μ'. Values of α for selected values of σ/μ are listed in Table 2.3.

Expressions for the skewness and kurtosis are obtained from Equation (2.21) for the appropriate values of r. Simplified as far as seems useful, the results are

$$\eta_{3l\text{-}l} = \frac{3\left(s^2+1\right)^2}{s^3\left[4-\sec^2(\pi/\alpha)\right]} - \frac{3s^2+1}{s^3}, \tag{2.24}$$

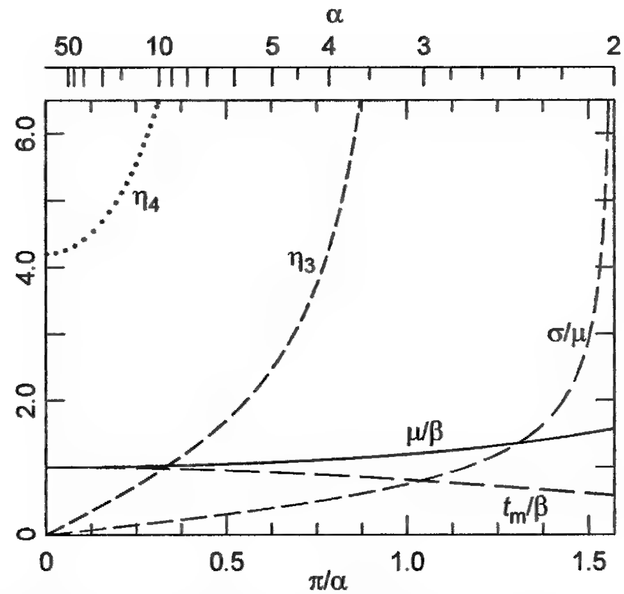

Figure 2.3: Various properties of the log-logistic distribution as a function of the parameter α for the case $\alpha > 2$. The curves show the ratios t_m/β, μ/β and σ/μ, where t_m is the mode, μ the mean and σ the standard deviation, and values of the skewness η_3 and kurtosis η_4.

Table 2.3: Values of α to six significant figures for various values of σ/μ.

σ/μ	α	σ/μ	α	σ/μ	α	σ/μ	α	σ/μ	α
0.05	36.3304	0.30	6.36448	0.70	3.26802	1.10	2.58774	3.0	2.08824
0.10	18.2465	1/3	5.79329	0.75	3.13310	1.20	2.50287	4.0	2.05007
0.15	12.2542	0.40	4.95149	0.80	3.01822	1.40	2.37947	5.0	2.03218
0.20	9.28422	0.50	4.13744	0.90	2.83440	1.60	2.29595	6.0	2.02240
0.25	7.52257	0.60	3.61953	1.00	2.69535	2.00	2.19380	10.0	2.00809

$$\eta_{4\text{l-l}} = \frac{(s^2+1)^3}{s^4\left[2-\sec^2(\pi/\alpha)\right]} - \frac{12(s^2+1)^2}{s^4\left[4-\sec^2(\pi/\alpha)\right]} + \frac{6s^2+3}{s^4}. \tag{2.25}$$

These are also plotted in Figure 2.3. As this Figure shows, the higher moments of the log-logistic distribution are rather large compared with distributions considered so far. Equations (2.23)–(2.25) also show that the parameter β is simply a scaling parameter.

Figure 2.3 indicates that the full range of values of σ/μ are accessible only if infinite higher moments are tolerated. If one wishes to consider only distributions with finite skewness or kurtosis, then this introduces an upper limit on σ/μ: $\sigma/\mu < 0.80869$ for finite skewness and $\sigma/\mu < 0.52272$ for finite kurtosis.

2.2.8 Inverted Gamma Distribution

The inverted gamma distribution is the distribution of a random variate whose reciprocal is gamma-distributed. The standard form of the distribution is [12(§18.4)]

$$f_{\text{i-}\Gamma}(t) = \frac{\lambda^c}{\Gamma(c)t^{c+1}}\exp\left(\frac{-\lambda}{t}\right) \qquad (0 \le t < \infty), \tag{2.26}$$

where the parameters λ, c must both be positive. Law and Kelton refer to this distribution as 'the' Pearson type V distribution [7], but it is actually just one example of this class of distributions [10(§12.4),16(ch.4)].

For this distribution, the requirement $f_{\text{i-}\Gamma}(0) = 0$ imposes no additional constraint on the parameters; however constraints arise when moments are calculated: for the rth moment to be finite, one requires $c > r$, as with the parameter α of the log-logistic distribution. Expressions are much simpler in the case of the inverted gamma distribution, so that closed-form expressions can be obtained with μ, σ as parameters. The transformation is

$$\lambda = \mu(\nu^2 + 1), \qquad c = \nu^2 + 2, \tag{2.27}$$

where $\nu = \mu/\sigma$, the reciprocal of the coefficient of variation. The resulting expressions for mode, skewness and kurtosis are listed in Table 2.1 (p. 7). The probability density is

$$f_{\text{i-}\Gamma}(t) = \frac{\left[\mu\left(\nu^2+1\right)\right]^{\nu^2+2}}{\Gamma\left(\nu^2+2\right) t^{\nu^2+3}} \exp\left(-\mu\frac{\nu^2+1}{t}\right). \tag{2.28}$$

As with the log-logistic distribution, the whole range of values of σ/μ is accessible provided that infinite higher moments are tolerated. If one wishes to consider only distributions with finite higher moments, then one requires $\sigma/\mu < 1$ to keep the skewness finite and $\sigma/\mu < 1/\sqrt{2}$ to keep the kurtosis finite.

2.2.9 Beta-Prime or Generalised *F* Distribution

The β' distribution is derived from the β distribution (§3.2 below) by a transformation of the argument. It is the only three-parameter distribution on $[0,\infty)$ considered in this report. The standard form is [7,10(§25.7,§27.8.1)][(b)]

$$f_{\beta'}(t) = \frac{\tau^q}{B(p,q)} \frac{t^{p-1}}{(t+\tau)^{p+q}} \qquad (0 \le t < \infty), \tag{2.29}$$

where all three parameters p, q, τ must be positive and $B(p,q)$ is the standard beta function [8(§6.2)]:

$$B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}. \tag{2.30}$$

The requirement $f_{\beta'}(0) = 0$ imposes the constraint $p > 1$; further constraints arise from the moments: a finite rth moment is obtained only if $q > r$. Equation (2.29) is a generalisation of the much studied *F* distribution of variance ratios [10(ch.27)].[(c)]

With three free parameters, one can specify the skewness as well as mean and standard deviation. An expression for the rth moment μ_r about the origin can be readily derived:

$$\mu_r = \tau^r \frac{(p+r-1)!\,(q-r-1)!}{(p-1)!\,(q-1)!}. \tag{2.31}$$

Working from this, the formulae for the three moments in terms of the parameters are:

(b) As with the inverted gamma distribution, Law and Kelton [7] misname this distribution, calling it 'the' Pearson type VI distribution, whereas it is but one example of the class [10 (§12.4,§27.7),16(ch.4)]. Johnson *et al.* [10(p.345)] present equations for the inversion of a 4-parameter generalisation of this distribution.

(c) The *F* distribution is obtained by setting $\tau = q/p$, i.e. it has two parameters only.

$$\mu = \frac{\tau p}{q-1}, \qquad \sigma^2 = \frac{\tau^2 p(p+q-1)}{(q-1)^2(q-2)}, \tag{2.32}$$

$$\eta_3 = 2\frac{2p+q-1}{sp(q-3)}, \tag{2.33}$$

where $s = \sigma/\mu$, as before. Equation (2.33) is independent of the parameter τ, and so is identical to the expression for the skewness of the F distribution, since the F distribution differs from the β' distribution only in its τ values. Equations (2.32) yield a simple result for s:

$$s^2 = \frac{p+q-1}{p(q-2)}. \tag{2.34}$$

Recalling that this expression assumes $q > 2$ and that we are also taking $p > 1$, we see that Equation (2.34) implies $0 < \sigma/\mu < \infty$; that is, the full range of s values is available. However, the inclusion of Equation (2.33) implies a further restriction of the range of q to $q > 3$, which then restricts the range of σ/μ to $0 < \sigma/\mu < \sqrt{3}$. The same argument applied to Equation (2.33) shows that η_3 must be positive; that is, negative skewness is not accessible for this distribution.

Equations (2.32) and (2.33) are sufficiently simple for the inversion to be relatively straightforward:

$$p = 2\frac{1+s\eta_3 - s^2}{s^3\eta_3 - s\eta_3 + 4s^2}, \qquad q = 3 + 2\frac{s^2+1}{s\eta_3 - 2s^2}, \tag{2.35}$$

with τ being calculable from the expression for μ once p and q are known. The requirements $p > 1$, $q > 3$ put restrictions on values of η_3 that may be validly chosen:

$$\eta_3 > \max\left(2s, \frac{6s^2-2}{3s-s^3}\right). \tag{2.36}$$

This analysis leaves the kurtosis η_4 as the first moment available for comparison with other distributions. Derivation shows that the expression for η_4 is also independent of the parameter τ and so is the same expression as for the F distribution; that is

$$\eta_4 = 3 + 6\frac{5q-11}{(q-3)(q-4)} + 6\frac{(q-1)^2(q-2)}{p(p+q-1)(q-3)(q-4)} \tag{2.37}$$

where it is assumed that $q > 4$. (The kurtosis is infinite for $q \le 4$.) This further restriction on q carries with it the implications

$$0 < s < \sqrt{2}, \qquad \eta_3 < 4s + 2/s. \tag{2.38}$$

2.3 Comparisons and Comments

2.3.1 General Distribution Properties

Table 2.1 (pp. 6–7) comprises a comparison of sorts between the distributions. This Section makes a more graphic comparison through plots of the distributions and their properties. From these plots and Table 2.1, several specific points of comparison are noted.

Figure 2.4 shows examples of distributions with the same mean and standard deviation. Comparison between the distributions is made through three properties: the

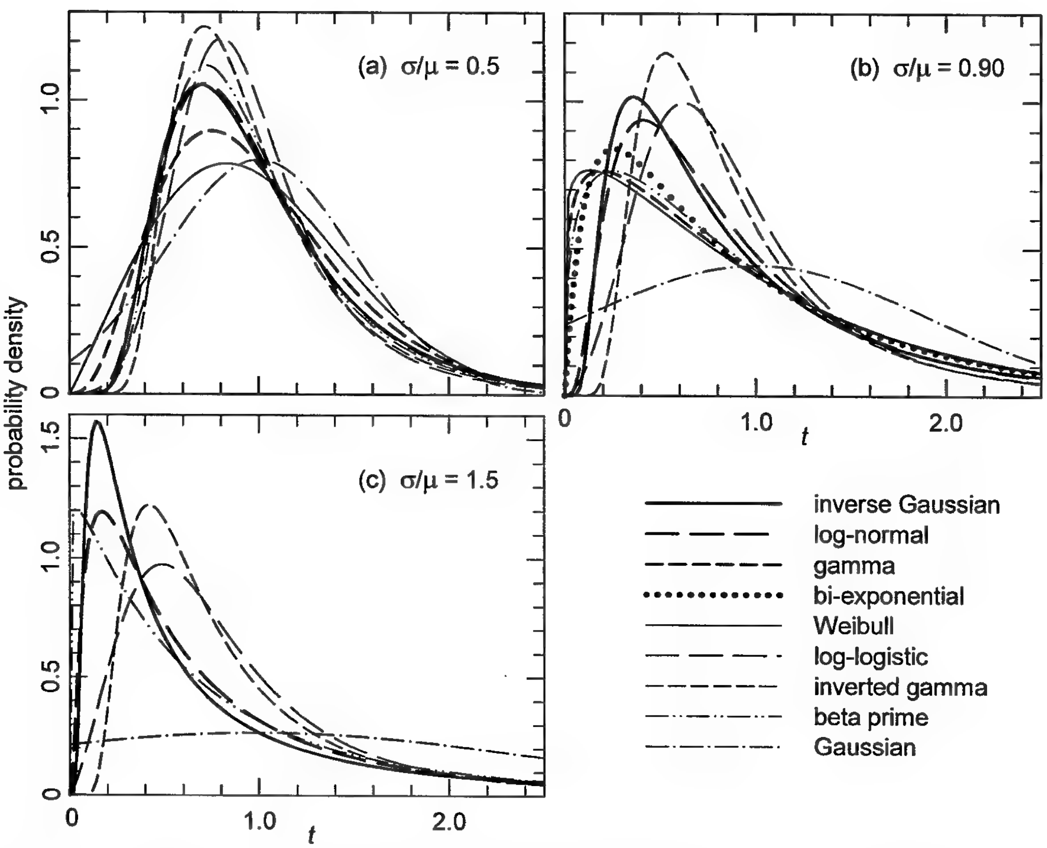

Figure 2.4: Examples of nine continuous distributions with mean μ of 1.0 and standard deviation σ of (a) 0.50, (b) 0.90 and (c) 1.50. The bi-exponential distribution is omitted from panel (a) because it cannot have this combination of mean and standard deviation (§2.2.5); several distributions are omitted from panel (c) for the same reason. The skewness values chosen for the β' distribution are (a,b) 2.0, (c) 12.0.

ratio t_m/μ of mode to mean, the skewness η_3 and the kurtosis η_4; values of these quantities are listed in Table 2.4. In view of their importance, Figure 2.5 compares these properties in a consistent manner by plotting their values as a function of σ/μ.[d] The β' distribution is different from the others in having a third parameter, which is taken in §2.2.9 as being given by the skewness. The range of valid skewness values and corresponding ranges of t_m/μ and η_4 are shown as shaded areas in Figure 2.5.

Figure 2.4 indicates the extent to which the distributions resemble one another, particularly at low σ/μ. As Figure 2.5 shows, the properties of all except the Weibull, log-logistic and β' distributions approach those of the Gaussian as $\sigma/\mu \to 0$. However, the following points of contrast and difference are noted:

- The inverse Gaussian, log-normal and log-logistic distributions have quite similar behaviour near $t = 0$ (Fig. 2.4). Compared with these, the inverted gamma distribution approaches zero more quickly and all the rest less quickly as $t \to 0$.

[d] For the Weibull and log-logistic distributions, Equations (2.16) and (2.23) cannot be inverted to obtain closed-form equations for the parameters as a function of σ/μ, but it is clearly possible to construct plots of the distribution properties against σ/μ from the data in Figures 2.2 and 2.3.

*Table 2.4: Values of the ratio t_m/μ of mean to mode, skewness η_3 and kurtosis η_4 for the distributions shown in Figure 2.4.**

	$\sigma/\mu = 0.5$			$\sigma/\mu = 0.9$			$\sigma/\mu = 1.5$		
	t_m/μ	η_3	η_4	t_m/μ	η_3	η_4	t_m/μ	η_3	η_4
inverse Gaussian	0.693	1.50	6.75	0.359	2.70	15.2	0.145	4.50	36.8
log-normal	0.716	1.63	8.04	0.411	3.43	29.4	0.171	7.87	209
gamma	0.750	1.00	4.50	0.190	1.80	7.9	-	-	-
bi-exponential	-	-	-	0.257	1.96	8.8	-	-	-
Weibull	0.830	0.57	3.13	0.121	1.70	7.2	-	-	-
log-logistic	0.805	3.86	244.0	0.623	∞	∞	0.489	∞	∞
inverted gamma	0.714	2.67	22.0	0.528	18.9	∞	0.419	∞	∞
beta prime	0.722	2†	12.0	0.249	2†	9.6	0.033	12†	∞
Gaussian	1.000	0	3.0	1.000	0	3.0	1.000	0	3

* A dash indicates that the value of σ/μ is not valid for that distribution.
† Chosen parameter value.

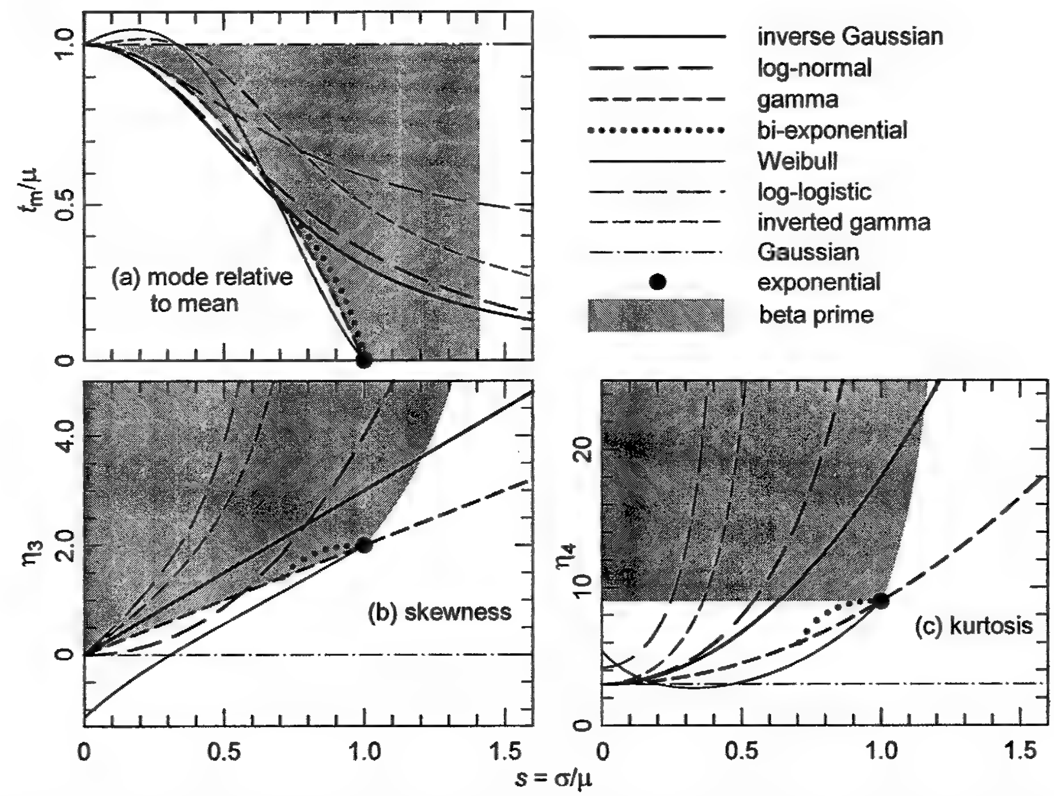

Figure 2.5: Comparison of several properties of the continuous distributions: (a) ratio t_m/μ of mode to mean, (b) skewness η_3 and (c) kurtosis η_4 are shown as a function of σ/μ. Note that the gamma and Weibull distributions are limited to $0 \le \sigma/\mu \le 1$, and the bi-exponential distribution to $1/\sqrt{2} \le \sigma/\mu \le 1$. The exponential distribution has $\sigma/\mu = 1$ only. The β' distribution has three parameters, and so its properties have a range of values for any given value of σ/μ. Skewness is taken as the third parameter; the shaded region in panel (b) shows the range of η_3 values satisfying Equations (2.36) and (2.38). The shaded areas in the other panels indicate the corresponding ranges of values of t_m/μ and η_4, although this representation does not show the detailed correspondence between points in these ranges.

- This difference in behaviour is due to the behaviour of the derivatives of the distributions as $t = 0$. There is a clear distinction between those that have all derivatives zero at regardless of parameter values and those for which only low-order derivatives are zero (Table 2.1). Only the inverse Gaussian, log-normal and inverted gamma distributions fall into the first category.
- In general, the inverse Gaussian distribution is less peaked and falls less quickly as $t \to \infty$ than the log-logistic and inverted gamma distributions, with the opposite applying for all the other distributions (Fig 2.4).
- The Weibull distribution is unique in having regions of σ/μ where the mode lies at larger t values than the mean, the skewness is negative and the kurtosis is less than the value of 3.0 that applies to the Gaussian (Fig. 2.5). Of these three features, one is shared by the log-logistic distribution, for which $t_m/\mu > 1$ at small values of σ/μ.
- The Weibull distribution is also unique in that its skewness remains non-zero as $\sigma/\mu \to 0$ (Fig. 2.5c).
- As $\sigma/\mu \to 0$, the kurtosis of most distributions approaches 3.0, the value for the Gaussian distribution. The Weibull and log-logistic distributions are the two exceptions to this behaviour (Fig. 2.5d). The log-logistic distribution does not have a kurtosis as small as 3.0 for any values of its parameters.
- The log-logistic and inverted gamma distributions have markedly higher skewness and kurtosis values for a given value of σ/μ than the other distributions. The only exception to this statement is the β' distribution, for which the skewness can be chosen over a wide range as desired. The resulting values of kurtosis for the β' distribution also span a wide range.

2.3.2 Limits on Ranges of Coefficients of Variation

A significant outcome of the discussion in §2.2 is the identification of limits on the accessible values of coefficients of variation σ/μ. These limits all bear on the question of which distributions are applicable to a given application. Because of their importance, the results are collected in Table 2.5.

Table 2.5: Accessible ranges of coefficients of variation for the distributions considered in §2.2.

Distribution	range of $s = \sigma/\mu$
inverse Gaussian	no limit
Gaussian	no limit
exponential	$s = 1$
log-normal	no limit
gamma	$s < 1$
bi-exponential	$1/\sqrt{2} < s < 1$
Weibull	$s < 1$
log-logistic	no limit
log-logistic, finite skewness	$s < 0.8087$
log-logistic, finite kurtosis	$s < 0.5227$
inverted gamma	no limit
inv. gamma, finite skewness	$s < 1$
inv. gamma, finite kurtosis	$s < 1/\sqrt{2}$
beta-prime	no limit
beta-prime, finite skewness	$s < 1/\sqrt{3}$
beta-prime, finite kurtosis	$s < 1/\sqrt{2}$

2.3.3 Fifth and Higher Moments

A recent study examined the effect of varying the distribution of service times in the application of queueing theory to the modelling of maritime interception operations [2]. In the course of this study, it was observed that the exponential distribution gives results noticeably different from the other distributions studied, which included the inverse Gaussian, log-normal and gamma distributions, among others. One may seek an explanation for this in the behaviour of the higher moments. The distributions concerned are the inverse Gaussian and log-normal, since these were the only two included in Ref. 2 that can have $\sigma = \mu$, a property characteristic of the exponential distribution. The values of skewness and kurtosis of the three distributions (Table 2.1, p. 6; Figs 2.5c,d) are not different enough to explain the effects observed in Ref. 2, so it is interesting to look at yet higher moments. It turns out that general expressions for these are available or can be readily derived.

The quantities of interest are the rth moments μ_r about the mean and the rth moment ratios η_r:

$$\mu_r = \int_0^\infty (t-\mu)^r f(t)\mathrm{d}t\,, \qquad \eta_r = \mu_r/\sigma^r, \tag{2.39}$$

the second being the natural generalisation of skewness and kurtosis. These quantities can be readily evaluated for the exponential distribution, giving

$$\eta_{re} = \sum_{k=0}^{r-2} (-1)^{r-k}\, k! \binom{r}{k} \qquad (r \geq 3), \tag{2.40}$$

where

$$\binom{r}{k} = \frac{r!}{k!\,(r-k)!} \tag{2.41}$$

is the binomial symbol.

For the log-normal distribution, Johnson *et al.* [10(§14.3)] give a formula the rth moment about the mean. Setting $\sigma = \mu$ in this expression, one finds

$$\eta_{r\text{l-n}} = \sum_{k=0}^{r} (-1)^k \binom{r}{k} 2^{(r-k)(r-k-1)/2} \qquad (r \geq 3). \tag{2.42}$$

For the inverse Gaussian distribution, Evans *et al.* [12] give an expression for the rth moment about the origin. This can be converted to the rth moment about the mean by a standard transformation [12]. Once again setting $\sigma = \mu$, one obtains

$$\eta_{r\text{iG}} = (-1)^{r-1}(r-1) + \sum_{k=0}^{r-2} (-1)^k \binom{r}{k} \sum_{l=0}^{r-k-1} \frac{(r-k+l-1)!}{(r-k-l-1)!\,l!\,2^l} \qquad (r \geq 3). \tag{2.43}$$

These three quantities are compared in Figure 2.6 up to $r = 10$. The impetus for this comparison was a study in which use of the exponential distribution gave results significantly different from the other two [2]. However, if one distribution can be said to behave differently from the others in Figure 2.6, it is the log-normal rather than the exponential. Clearly, the explanation of the observed effect must be sought elsewhere. One possibility concerns modes (Fig. 2.5a): the mode of the exponential distribution is at $t = 0$; for the other two it lies at $0.303\,\mu$ and $0.354\,\mu$ (inverse Gaussian and log-normal respectively with $\sigma = \mu$). Another possibility is the decision rates, discussed in the next Section.

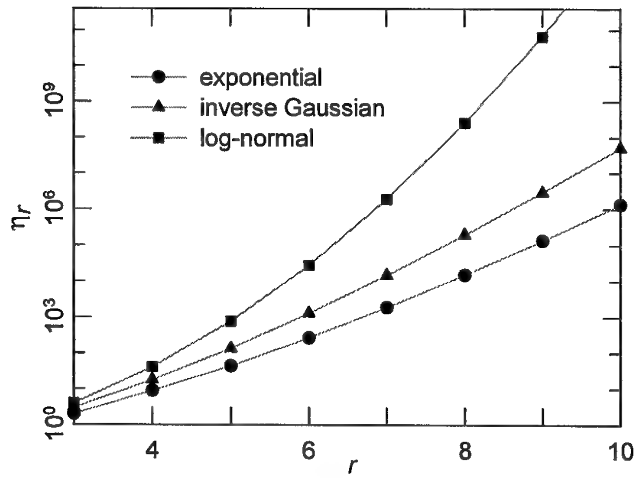

Figure 2.6: Values of the r*th moment ratios* η_r *of the exponential distribution and the inverse Gaussian and log-normal distributions with* $\sigma = \mu$*. Note the logarithmic scale on the ordinate. The points at* $r = 3$ *and* $r = 4$ *are the skewness and kurtosis respectively. The lines are intended solely as guides to the eye.*

2.3.4 Decision Rate

In the context of distributions of decision times, the *decision rate* $h(t)$ is the probability density of a decision occurring at time t on the condition that a decision has not occurred at times earlier than t. That is, it is given by

$$h(t) = \frac{f(t)}{1 - F(t)} \tag{2.44}$$

where $F(t)$ is the cumulative probability corresponding to the probability density $f(t)$: (e)

$$F(t) = \int_0^t f(t')\mathrm{d}t' . \tag{2.45}$$

The quantity defined in Equation (2.44) has also been termed 'hazard rate' [3,12,13,17], 'failure rate' [12,17–19] and 'force of mortality' [12,13]. It is widely used as a means of classifying distributions (e.g. [10(§33.2),19]).

Many of the probability distributions considered here have surprisingly simple expressions for their decision rate, albeit in terms of a variety of special functions. The inverse Gaussian is not one such; its decision rate is [18]

$$h_{\mathrm{iG}}(t) = \frac{f_{\mathrm{IG}}(t)}{\Phi\left[\nu(1 - t/\mu)\sqrt{\mu/t}\right] - e^{2\nu^2}\Phi\left[-\nu(1 + t/\mu)\sqrt{\mu/t}\right]}, \tag{2.46}$$

where $\Phi(t)$ is the cumulative probability of the Gaussian distribution with mean of zero and unit standard deviation.

The expressions for the decision rate, simple or otherwise, are collected in Table 2.1 (pp. 6–7). The point leading to an argument for using the inverse Gaussian distribution to modelling decision speed concerns the behaviour of $h(t)$ over its domain and particularly as $t \to \infty$. Surveying the various distributions, we see that:

(e) The lower limit of the integral must be taken as $-\infty$ for two-sided distributions such as the Gaussian.

- The exponential distribution has a constant decision rate equal to $1/\mu$.
- All distributions other than the exponential and Gaussian have $h(0) = 0$.(f)
- The Weibull distribution with $\beta > 1$ (the case of interest in this Report) and β' distribution have monotonically increasing decision rates with no upper limit.
- The bi-exponential distribution and gamma distribution with $r > 1$ also have monotonically rising decision rates, but they asymptote as $t \to \infty$ to $h_{b\text{-}e} = \alpha$ and $h_\Gamma = \theta = \mu/\sigma^2$.

Except for the property $h(0) = 0$, none of this behaviour seems realistic in a model of decision rate. As the following dot points explain, it is likely that real-life decision rates rise from zero at $t = 0$ to a maximum value and then decrease to a non-zero value as $t \to \infty$:

- The rate should be low initially while the decision maker assesses the situation.
- It would then rise as the decision-maker's situation awareness increases and recognition-primed decisions [1] are made.
- This would be followed by a decline as the 'easy' decisions are taken, leaving the more difficult ones to be dealt with.
- However, the decision rate ought not to decline to zero; while ever a decision remains to be made, one expects a non-zero probability that it will be made.

Of the distributions in Table 2.1, *only* the inverse Gaussian distribution has a decision rate that rises from zero, peaks and then declines to a non-zero value as $t \to \infty$. That is, only the inverse Gaussian distribution fulfils all of the requirements. This is somewhat remarkable, in view of the diversity of distributions compiled in Table 2.1.

Figure 2.7 shows the time dependence of the decision rate for three distributions with $\sigma = \mu$, being the three distributions discussed in the previous Section. Decision rate provides another aspect that distinguishes the exponential distribution from the other two.

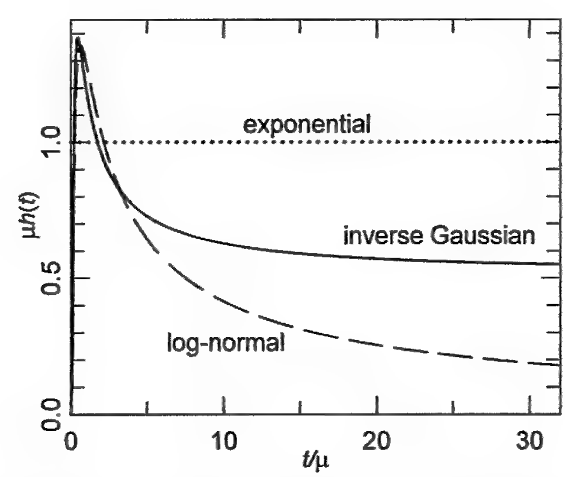

Figure 2.7: Time dependence of decision rates $h(t)$ *for three distributions with* $\sigma = \mu$. *To exploit scaling behaviour,* $\mu h(t)$ *is plotted against* t/μ. *The product* $\mu h(t)$ *for the log-normal distribution asymptotes to zero as* $t \to \infty$; *that of the inverse Gaussian distribution asymptotes to 0.5 (in the case* $\sigma = \mu$).

(f) For the gamma and Weibull distributions, this is true only for parameter values of interest in this report, i.e. those giving $s < 1$. (When $s = 1$, both distributions degenerate to the exponential distribution and, for $s > 1$, both start at $h(0) = \infty$ and fall monotonically as $t \to \infty$.)

3. Finite-Domain Distributions

The initial interest in distributions on a finite domain arose from considerations of modelling decision soundness, since soundness scales typically are finite. If the scale is continuous, the β distribution is useful because of the wide range of distribution shape available. Often, soundness scales are discrete ('Likert scales'), so there is interest in discrete probability distributions from that point of view. Both these topics are discussed in this Section, but first the following observation is addressed.

Table 2.1 and Figure 2.5(c) show that the Gaussian distribution has a kurtosis of 3, independent of the value chosen for the standard deviation, and that almost all other distributions have larger kurtosis. The sole exception is the Weibull distribution for a limited range of values of β, and even there the minimum value available is ~2.71, not much less than 3. In the interests of exploring a wide range of behaviour, it may be useful on some occasions to use distributions with markedly lower kurtosis than that of the Gaussian, even if their other properties make them unlikely candidates for, say, a realistic description of decision times.

3.1 Simple Low-Kurtosis Continuous Distributions

This Section presents properties of three simple distributions, two of which clearly have rather lower kurtosis that the Gaussian distribution. Their definitions are given below and their properties are collected in Table 3.1.

3.1.1 Impulse Distribution

The impulse or deterministic distribution has a single parameter, its mean μ, and is defined by

$$f_{\mathrm{i}}(t) = \delta(t-\mu), \tag{3.1}$$

where $\delta(x)$ is the impulse symbol, also known as the Dirac delta function. The effect of Equation (3.1) is that μ is the only possible value for t; the probability of obtaining any other value is strictly zero. The corresponding cumulative probability (Eqn (2.45)) is the unit step function located at $t = \mu$.

Table 3.1: Properties of three simple probability density functions with finite domain.

	Impulse	Uniform	Triangular
Range of $s = \sigma/\mu$	$s = 0$	$s < 1/\sqrt{3}$	$s < 1/\sqrt{8}$ to $1/\sqrt{2}$*
Mode t_m	μ	none	t_m
Skewness η_3	0	0	Eqn (3.7)
Kurtosis η_4	undefined†	9/5	12/5
Decision rate $h(t)$‡	$\delta(t-\mu)$	$\begin{cases} \dfrac{1}{b-t} & (a<t<b) \\ 0 & (\text{otherwise}) \end{cases}$	$\begin{cases} \dfrac{2(t-a)}{(b-a)t_m - t^2 + 2at - ab} & (a \le t \le t_m) \\ \dfrac{2}{b-t} & (t_m \le t \le b) \\ 0 & (\text{otherwise}) \end{cases}$

* Depending on η_3, see §3.1.3. † See §3.1.1. ‡ See §2.3.4.

It may seem that this distribution has the smallest possible kurtosis, namely zero, since all moments about the mean are zero. However, the kurtosis is defined as the ratio of two such moments, the fourth to the square of the second, and it turns out that the result is strictly undefined. To see this, consider the common method of introducing the impulse symbol using a limiting sequence of appropriate functions [e.g. 20(ch. 5)]. Choosing a sequence of rectangles gives a kurtosis of 9/5; a sequence of triangles gives a value of 12/5; one of Gaussians gives 3, etc. Since all these sequences are held to represent exactly the same object, the impulse symbol, it follows that its kurtosis is not a well-defined quantity.(g) Nevertheless, there is no doubt that the impulse distribution represents an extreme of behaviour among probability density functions, and this in itself makes it useful in some circumstances.

3.1.2 Uniform distribution

The uniform distribution, perhaps the simplest of all distributions, has a kurtosis of just 9/5, the lowest encountered in this work among the continuous distributions.(h) The probability density is

$$f_{\mathrm{u}}(t) = \begin{cases} \dfrac{1}{b-a} & (a \le t \le b) \\ 0 & (\text{otherwise}) \end{cases} \tag{3.2}$$

where we require $a > 0$ to ensure that $f_{\mathrm{u}}(0) = 0$. The mean and standard deviation are

$$\mu = \frac{a+b}{2}, \qquad \sigma = \frac{b-a}{\sqrt{12}}. \tag{3.3}$$

Rewritten using σ, μ as parameters, Equation (3.2) is

$$f_{\mathrm{u}}(t) = \begin{cases} \dfrac{1}{2\sigma\sqrt{3}} & (\mu - \sigma\sqrt{3} \le t \le \mu + \sigma\sqrt{3}) \\ 0 & (\text{otherwise}). \end{cases} \tag{3.4}$$

Hence, the requirement $a > 0$ places a limitation on the values of $s = \sigma/\mu$ that can be obtained: $s < 1/\sqrt{3}$. The uniform distribution does not have a mode—it has no peak. Other properties are listed in Table 3.1.

3.1.3 Triangular Distribution

The triangular distribution has probability density that is zero for $t \le a$, rises linearly from $t = a$ to the mode t_{m}, falls linearly from $t = t_{\mathrm{m}}$ to zero at $t = b$ and is zero thereafter:

$$f_{\mathrm{t}}(t) = \begin{cases} \dfrac{2(t-a)}{(b-a)(t_{\mathrm{m}}-a)} & (a \le t \le t_{\mathrm{m}}) \\ \dfrac{2(t-a)}{(b-a)(t_{\mathrm{m}}-a)} & (t_{\mathrm{m}} \le t \le b) \\ 0 & (\text{otherwise}), \end{cases} \tag{3.5}$$

(g) The skewness of the impulse distribution is well defined. All valid limiting sequences must be sequences of functions that are symmetric about μ, and hence all agree that the skewness is zero.

(h) Only the discrete uniform (§3.3.1) and β-binomial distributions (§§3.3.4 and 3.3.5) have lower accessible kurtosis values.

where a ≥ 0 to keep $f_t(0) = 0$. Equation (3.5) is a generalisation of the frequently anthologised symmetric triangular distribution, which is obtained by setting $t_m = (a + b)/2$. The mean and standard deviation of the general distribution are listed in tabulations [7,12], but the higher moments seem not to have been previously published. The mean and variance are

$$\mu_t = \frac{a + t_m + b}{3}, \qquad \sigma_t^2 = \frac{a^2 + t_m^2 + b^2 - at_m - ab - bt_m}{18}. \tag{3.6}$$

Straightforward, though lengthy, calculation gives the skewness as

$$\eta_{3t} = \frac{2\left(a^3 + t_m^3 + b^3\right) - 3\left(a^2 t_m + a t_m^2 + a^2 b + ab^2 + b^2 t_m + b t_m^2\right) + 12abt_m}{270\sigma^3}. \tag{3.7}$$

The kurtosis is $\eta_{4t} = 12/5$ *regardless of the parameter values*. The η_{4t} value is well known for the symmetric triangular distribution, but it does not seem to have been previously recognised that it applies in general.(i) It is a most remarkable result. It means that the skewness can be varied while holding mean, standard deviation *and* kurtosis constant, despite the fact that the triangular distribution has just three parameters. This property seems to be unique among probability distributions. It recently proved useful in a study of the effect on a queueing system of varying service-time distributions [2].

The inversion of Equations (3.6) and (3.7) is algebraically challenging, but was performed with the aid of a computer algebra package. The result is

$$\begin{aligned} a &= \mu_t - \sigma_t\sqrt{2}\cos\xi - \sigma_t\sqrt{6}\sin\xi, \\ t_m &= \mu_t - \sigma_t\sqrt{2}\cos\xi + \sigma_t\sqrt{6}\sin\xi, \\ b &= \mu_t + 2\sigma_t\sqrt{2}\cos\xi, \end{aligned} \tag{3.8}$$

where the auxiliary angle ξ is given by

$$\xi = \frac{1}{3}\arctan\frac{\sqrt{8 - 25\eta_{3t}^2}}{5\eta_{3t}}. \tag{3.9}$$

To evaluate Equation (3.9) appropriately, the result of the arctan function must be taken in the range $[0,\pi]$, which is not the usual range assumed by calculators and software routines. That is, $\eta_{3t} = 0$ corresponds to $\xi = \pi/6$ and negative values of η_{3t} correspond to $\pi/6 < \xi \le \pi/3$.

In the course of the derivation of Equations (3.8) and (3.9), one finds that the available range of skewness values is limited to

$$|\eta_{3t}| \le \frac{\sqrt{8}}{5}; \tag{3.10}$$

unlike the case of the β' distribution, negative skewness is accessible. The limits correspond to the 'sawtooth distributions': $\eta_{3t} = -\sqrt{8}/5$ ($\xi = \pi/3$) when $t_m = b$ and $\eta_{3t} = +\sqrt{8}/5$ ($\xi = 0$) when $t_m = a$. The first line of Equation (3.8) implies a limit on the value of σ from the requirement that $a \ge 0$:

(i) Johnson *et al.* [10(§26.9)] quote a formula for general moments about t_m. Ayyangar [21] gives an expression for moments about the mean in terms of moments about t_m, but evaluates it for the symmetric distribution only.

$$s_t = \frac{\sigma_t}{\mu_t} \le \frac{\operatorname{cosec}(\xi + \pi/6)}{\sqrt{8}}. \tag{3.11}$$

That is, the value of the limit depends on the chosen skewness: the limit is quite low for maximum negative skewness ($s_t \le 1/\sqrt{8}$ when $\eta_{3t} = -\sqrt{8}/5$) and rises as η_{3t} is increased, reaching $s_t \le 1/\sqrt{2}$ when $\eta_{3t} = +\sqrt{8}/5$.

3.2 The General Beta Distribution

The general β distribution is a very general finite-domain continuous distribution. Having four parameters, a wide range of curve shapes is encompassed. Its standard form is [10(ch.25)]

$$f_\beta(t) = \begin{cases} \dfrac{(t-a)^{p-1}(b-t)^{q-1}}{B(p,q)(b-a)^{p+q-1}} & (a \le t \le b) \\ 0 & \text{(otherwise)}, \end{cases} \tag{3.12}$$

where $B(p,q)$ is the beta function (Eqn (2.30)) and p, q must both be positive. We consider only those distributions that are finite at both end points, which implies $p \ge 1$, $q \ge 1$. In addition, we require $a \ge 0$, as usual. The case $p = 1$, $q = 1$ is exactly the uniform distribution on $[a,b]$, and the cases $p = 1$, $q = 2$ and $p = 2$, $q = 1$ are the two sawtooth distributions mentioned in §3.1.3. Obviously, the kurtosis of the β distribution is low for at least part of its parameter space.

The properties of the β distribution are well known [10(ch.25),13(ch.14)]. The mean μ and variance σ^2 are

$$\mu = a + \frac{(b-a)p}{p+q}, \qquad \sigma^2 = \frac{(b-a)^2 pq}{(p+q)^2(p+q+1)}; \tag{3.13}$$

the skewness η_3 and kurtosis η_4 are

$$\eta_3 = \frac{2(q-p)}{p+q+2}\sqrt{\frac{p+q+1}{pq}}, \qquad \eta_4 = \frac{3(p+q+1)\left[2(p-q)^2 + pq(p+q+2)\right]}{pq(p+q+2)(p+q+3)}; \tag{3.14}$$

and, for the case $p > 1$, $q > 1$, the mode t_m is

$$t_m = a + \frac{(b-a)(p-1)}{p+q-2}. \tag{3.15}$$

Following the theme of this Report, we wish to invert Equations (3.13) and (3.14) to obtain expressions for a, b, p, q in terms of μ, σ, η_3, η_4. Of the distributions included in this Report,(j) the β distribution is the only one where such expressions(k) have been found elsewhere [10(§25.4))]. In terms of the auxiliary variables

---

(j) See also footnote (b) (p. 11).

(k) As Johnson *et al.* note [10(§25.4)], sometimes values of a, b are set by external considerations and the requirement is to choose p, q to give desired values of μ, σ. This is a relatively simple problem, the solution to which is

$$p = \frac{\mu - a}{b-a}\left[\frac{(\mu-a)(b-\mu)}{\sigma^2} - 1\right], \qquad q = \frac{b-\mu}{b-a}\left[\frac{(\mu-a)(b-\mu)}{\sigma^2} - 1\right].$$

$$r = 6\frac{\eta_4 - \eta_3^2 - 1}{3\eta_3^2 - 2\eta_4 + 6}, \qquad R = \frac{(r+2)\eta_3}{\sqrt{(r+2)^2\eta_3^2 + 16(r+1)}}, \tag{3.16}$$

the expressions for p, q are

$$p = r(1-R)/2, \qquad q = r(1+R)/2. \tag{3.17}$$

A third auxiliary variable facilitates expressions for a, b:

$$A = \frac{\sigma}{2}\sqrt{(r+2)^2\eta_4 + 16(r+1)}; \qquad a = \mu - \frac{A(1-R)}{2}, \qquad b = A + a. \tag{3.18}$$

What is missing from the discussion of Ref. 10 is a description of the accessible range of skewness, kurtosis and coefficient-of-variation values. This turns out to be a complicated problem, with many factors potentially contributing to the boundaries of the accessible parameter space. For example, both A and R must be real, which means that η_3, η_4 are limited to values for which the square-root functions have non-negative arguments. Other requirements are $p \geq 1$ and $q \geq 1$. Figure 3.1 shows an accessible area in η_3–η_4 space. The upper boundary is part of the line of singularity of r; the lower boundary comes from the condition that $p \geq 1$ and $q \geq 1$. That is, the solid line in Figure 3.1 is given by

$$\eta_4 = 3 + 3\eta_3^2/2, \tag{3.19}$$

the condition for which the denominator of r is zero. The requirements $p \geq 1$ and $q \geq 1$ generate a cubic form in η_3^2 and η_4. Treated as an equation in η_4, this has three real roots, the largest of which for any given η_3^2 value is shown as the broken line in Figure 3.1. The equation of this line is

$$\eta_4 = M\cos\varphi + M\sqrt{3}\sin\varphi + \frac{60 + 54\eta_3^2 - 3\eta_3^4/8}{50 - \eta_3^2}, \tag{3.20}$$

where

$$M = \frac{3\sqrt{(\eta_3^2 + 100)(\eta_3^2 + 4)^3}}{8\left|50 - \eta_3^2\right|}, \qquad \varphi = \frac{1}{3}\arctan\frac{32\eta_3(\eta_3^2 - 50)}{2000 - 360\eta_3^2 + \eta_3^4}. \tag{3.21}$$

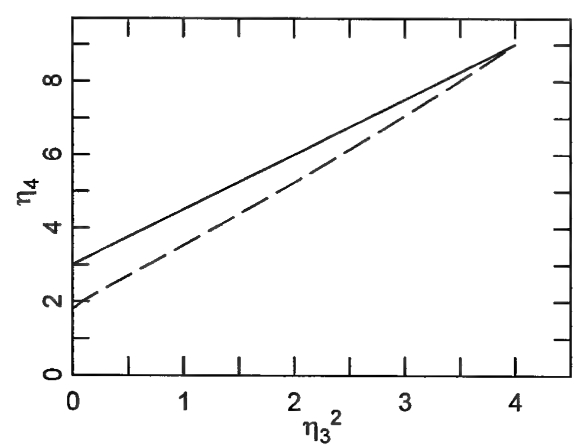

Figure 3.1: An area of accessible values of η_3 and η_4 for the general β distribution, shown as the region between the two lines. It is conjectured that this is the only such area.

Although not proven, it is conjectured on the basis of algebraic analysis and some trial-and-error evaluations that the region between the lines in Figure 3.1 is the only area of accessible η_3 and η_4 values for the general β distribution. The highest values are $|\eta_3| = 2$, $\eta_4 = 9$, like the exponential distribution. The lowest kurtosis accessible is $\eta_4 = 9/5$, which can only be achieved when $\eta_3 = 0$; in other words, this is exactly the uniform distribution.

The question of the range of σ/μ values accessible has not been explored, but it can be expected that the condition $a \geq 0$ places an upper limit on σ/μ, the value of which is a function of η_3 and η_4.

3.3 Discrete Probability Distributions

The use of a discrete-valued scale of measurement is common in the surveys used to gather data on decision soundness in specific situations. Hence, there is interest in discrete probability distributions for modelling decision soundness. Because the scales used often have few levels, for example 0–5 or 1–7, there is no necessity to employ a formula-based distribution: it would be feasible to regard a scale with $n + 1$ levels as having n parameters—namely the probabilities of obtaining each of n of the levels in any given trial—subject only to the restriction that the sum of the probabilities for all $n + 1$ levels must be unity. However, the degree of freedom entailed by this approach can be uncomfortably large, giving rise to a desire for a distribution with a small number of parameters. There is evidently no theoretical basis for choosing any particular probability distribution for the purpose of modelling decision soundness, so this Section compares the properties of several discrete probability distributions.

At first sight, there seem to be few candidates; most references list just three discrete distributions with finite domain: the uniform, binomial and hypergeometric. More comprehensive texts [6,12,22,23] mention a fourth, the β-binomial distribution, which is not, in fact, independent of the other three. It is unsurprising to learn that there are many more, but these are treated only in journals, specialised texts (e.g. [24]) and the most comprehensive compilations on discrete distributions [25,26]. Since these other distributions are less known, only the four distributions named above are treated here in any detail. Of these, the uniform distribution is of little interest, as explained in §3.3.1. The main properties of the other three are collected in Table 3.2.

Section 3.3.6 lists a few properties of some of the lesser known distributions, primarily as a cursory indication of the range of distributions that have been studied. For many of these, our program of expressing parameters in terms of moments cannot be carried through because either there are no closed-form expressions for the moments, or the expressions that exist are algebraically intractable.

3.3.1 Discrete Uniform Distribution

The discrete uniform distribution is[(l)]

$$p_{\mathrm{du}}(t) = \frac{1}{n+1} \qquad (0 \leq t \leq n). \tag{3.22}$$

Often, n is determined by external factors, such as the nature of the particular Likert scale used; in these cases, this distribution can be regarded as having no parameters.

[(l)] The lower limit of the domain is taken as zero rather than the more usual value of unity in the interests of compatibility with the other discrete distributions treated herein.

Table 3.2: Properties of three discrete probability density functions defined on $0 \le t \le n$.

	Binomial	Hypergeometric	Beta binomial
Mean μ	na	$\dfrac{nM}{N}$	$\dfrac{na}{a+b}$
Mode t_m*	$\lfloor (n+1)a \rfloor$	$\left\lfloor \dfrac{(n+1)(M+1)}{N+2} \right\rfloor$	$\left\lfloor \dfrac{(n+1)(a-1)}{a+b-2} \right\rfloor$
Variance σ^2	$na(1-a)$	$\dfrac{nM(N-n)(N-M)}{N^2(N-1)}$	$\dfrac{nab(a+b+n)}{(a+b)^2(a+b+1)}$
Skewness η_3	$\dfrac{1-2a}{\sigma_b}$	$\dfrac{(N-2n)(N-2M)}{N(N-2)\sigma_h}$	$\dfrac{(b-a)(a+b+2n)}{(a+b)(a+b+2)\sigma_{\beta\text{-}b}}$
Kurtosis η_4	$3+\dfrac{1}{\sigma_b{}^2}-\dfrac{6}{n}$	$\dfrac{N(N+1)-6n(N-n)}{(N-2)(N-3)\sigma_h{}^2}+\dfrac{3(N-1)(N+6)}{(N-2)(N-3)}-\dfrac{6M(N-M)}{(N-2)(N-3)\sigma_h{}^2}$	$\dfrac{a^2-4ab+b^2-a-b+3\mu b(a+b+n)}{(a+b+2)(a+b+3)\sigma_{\beta\text{-}b}{}^2}+\dfrac{6(a+b+1)\left(a^2-ab+b^2\right)}{ab(a+b+2)(a+b+3)}$

*$\lfloor x \rfloor$ = largest integer $\le x$. If x is exactly integral, then there are two equal maximum values, at $\lfloor x \rfloor$ and $\lfloor x \rfloor - 1$. If an expression exceeds n, then the mode is at n; if less than zero, then it is at zero.

This point of view is adopted here. Hence, the distribution is of little interest in the present context, since there is nothing to adjust. For reference, its moments are:

$$\mu_{du}=n/2, \qquad \sigma_{du}{}^2=n(n+2)/12, \qquad \eta_{3du}=0, \qquad \eta_{4du}=9/5-1\big/\left(5\sigma_{du}{}^2\right). \tag{3.23}$$

3.3.2 Binomial Distribution

The binomial distribution has a single parameter a satisfying $0 \le a \le 1$. The distribution, defined for integer t in the range $[0,n]$, is

$$p_b(t)=\binom{n}{t}a^t(1-a)^{n-t}, \tag{3.24}$$

where the first bracketed factor on the right-hand side is the binomial symbol (Eqn (2.41), p. 16). As Table 3.2 indicates, the parameter a equals μ_b/n. This leads to simple expressions for variance, skewness and kurtosis in terms of the mean; values are shown in Figure 3.2 for n = 10. As this Figure indicates, the full range of skewness, positive and negative, occurs. It is interesting that the minimum value of kurtosis, which occurs when $\mu_b = n/2$, equals $3 - 2/n$, a little less than the value of 3 applying to the Gaussian distribution.

3.3.3 Hypergeometric Distribution

The standard form of the hypergeometric distribution reads [6]

$$p_h(t)=\binom{M}{t}\binom{N-M}{n-t}\binom{N}{n}^{-1}, \tag{3.25}$$

where $0 \le t \le n$. The parameters M and N must both be integers and satisfy $M \ge n$, $N \ge M + n$. The requirement for M, N to be integers leads to a complication in the inversion

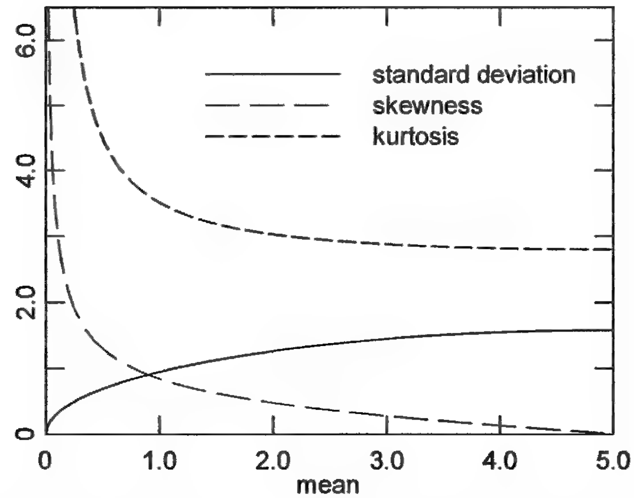

Figure 3.2: Standard deviation, skewness and kurtosis of the binomial distribution as functions of the mean for n = 10. *Only the left half of the domain is shown; the standard deviation and kurtosis are symmetric about* $\mu_b = 5$, *the skewness antisymmetric.*

of the expressions for μ, σ. The expressions (Table 3.2) can be readily inverted if *N, M* are treated as real numbers, called ***N, M*** in the following equations:

$$\boldsymbol{N} = n\frac{\mu(n-\mu)-\sigma^2}{\mu(n-\mu)-n\sigma^2}, \qquad \boldsymbol{M} = \mu\frac{\mu(n-\mu)-\sigma^2}{\mu(n-\mu)-n\sigma^2} \tag{3.26}$$

To complete the inversion, ***N, M*** must be converted to integers to obtain *N, M* respectively, but it is not clear how to formulate a rule for rounding that will lead to μ and σ values as close as possible to the desired values. Presumably each instance should be treated on its merits, using trial and error.

The question of the accessible ranges of μ, σ turns out to reveal an aspect of the connection between this, the binomial and the β-binomial distributions. Hence, it is deferred to §3.3.5, where the accessible ranges of skewness and kurtosis for the hypergeometric and β-binomial distributions are also determined.

3.3.4 Beta-Binomial Distribution

The β-binomial distribution is [6,12(p.37)]

$$p_{\beta\text{-b}}(t) = \binom{n}{t}\frac{B(a+t,b+n-t)}{B(a,b)}, \tag{3.27}$$

where once again $0 \le t \le n$. The parameters *a, b* need not be integers but both must be greater than zero. The function *B(a,b)* is the standard beta function (Eqn (2.30), p. 11), which is a generalisation of the binomial symbol to non-integer arguments. This distribution is also known as the negative or inverse hypergeometric distribution [22(pp. 155–60),23(pp.330–2),26(§6.2.2),27], a terminology that reflects the connection between the two distributions [26(§6.2.2),27].

Inversion of the equations for mean and variance (Table 3.2) gives results rather similar to Equation (3.26), another indication of the connections between the two distributions. Here, we have

$$a = \mu \frac{\mu(n-\mu)-\sigma^2}{n\sigma^2 - \mu(n-\mu)}, \qquad b = (n-\mu)\frac{\mu(n-\mu)-\sigma^2}{n\sigma^2 - \mu(n-\mu)}, \tag{3.28}$$

but this time the expressions can stand as they are, since a, b are not required to be integers. The question of the accessible ranges of μ and σ is again deferred to the next Section, which also deals with the ranges of skewness and kurtosis accessible to this distribution.

3.3.5 Comparisons and Ranges of Coefficient of Variation

Figure 3.3 illustrates the three distributions of interest. This figure differs from Figure 2.4 in that the latter shows distributions with the same mean and standard deviation. The distributions in Figure 3.3 all have the same mean, but different standard deviations. In fact, with finite parameter values it is not possible for any of the three discrete distributions to have simultaneously the same mean and standard deviation, as the following argument shows. If one expresses each standard deviation σ in Table 3.2 in terms of the corresponding mean μ, one obtains

$$\sigma_b^2 = \mu_b\left(1-\frac{\mu_b}{n}\right), \qquad \sigma_h^2 = \mu_h\left(1-\frac{\mu_h}{n}\right)\frac{N-n}{N-1},$$
$$\sigma_{\beta\text{-b}}^2 = \mu_{\beta\text{-b}}\left(1-\frac{\mu_{\beta\text{-b}}}{n}\right)\frac{a+b+n}{a+b+1}. \tag{3.29}$$

Now $n > 1$ in all cases and $N > n$ for the hypergeometric distribution, so it follows that distributions with equal means have $\sigma_h \le \sigma_b$, with the equality being approached as $N \to \infty$. On the other hand, the β-binomial distribution shows the opposite behaviour—$\sigma_{\beta\text{-b}} \ge \sigma_b$—since a and b are both positive; equality is approached as one of a or b goes

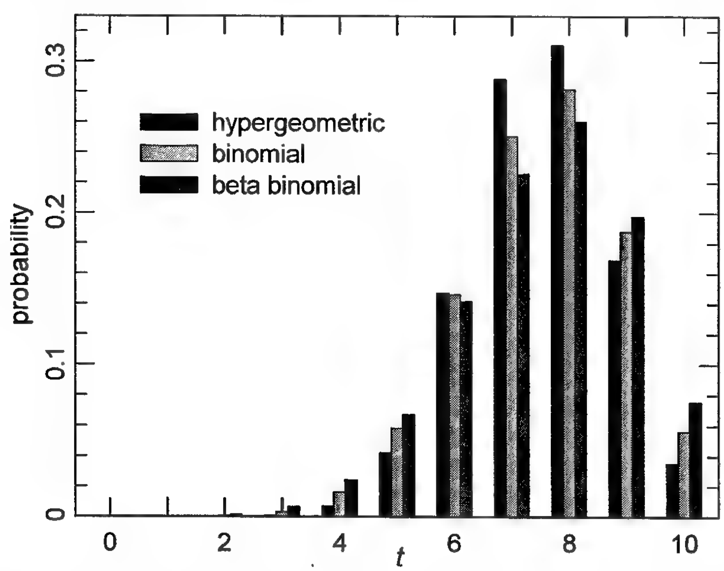

Figure 3.3: Discrete distributions with a maximum range of n = 10 and a mean of μ = 7.5. The standard deviation of the binomial distribution is fixed by the mean and range to be $\sqrt{1.875}$ = 1.369. The standard deviations of the other two distributions were chosen as σ_h = 1.20 and $\sigma_{\beta\text{-}b}$ = 1.50. (See text for discussion of these σ values.)

to infinity. In summary, distributions with equal means have their standard deviations ordered as $\sigma_h \le \sigma_b \le \sigma_{\beta\text{-}b}$, with strict inequalities for finite parameter values. Hence, the three discrete distributions are not alternatives in the sense that the various continuous distributions are; which is chosen depends on the mean and standard deviation that one wishes to model.

The other restrictions on the parameters lead to limits on the standard deviations attainable with the hypergeometric and β-binomial distributions that mirror those given above. In the case of the hypergeometric distribution, N has a lower limit determined by the values of μ_h and n. This gives a lower limit on the standard deviation:

$$\sigma_h \ge \mu_h \times \begin{cases} \dfrac{n-\mu_h}{\sqrt{\mu_h\left(n^2-\mu_h\right)}} & (\mu_h < n/2) \\ \sqrt{\dfrac{n-\mu_h}{n^2-n+\mu_h}} & (\mu_h \ge n/2). \end{cases} \tag{3.30}$$

For the β-binomial distribution, the parameter b can be regarded as determined by the values of a, $\mu_{\beta\text{-}b}$ and n. The lower limit on a ($a \ge 0$) then places an *upper* limit on the standard deviation:

$$\sigma_{\beta\text{-}b} < \mu_{\beta\text{-}b}\sqrt{n/\mu_{\beta\text{-}b} - 1} \tag{3.31}$$

The various limits on the standard deviations are depicted as a function of the mean in Figure 3.4 for the case of $n = 10$. In view of the relatively wide range of standard deviation accessible to the β-binomial distribution, it deserves to be better known. It should be noted that, although the region accessible to the hypergeometric distribution is shown as a structureless area in Figure 3.4 (the area with darker shading), in fact it consists of a lattice of isolated points, owing to the requirement that the parameters N, M

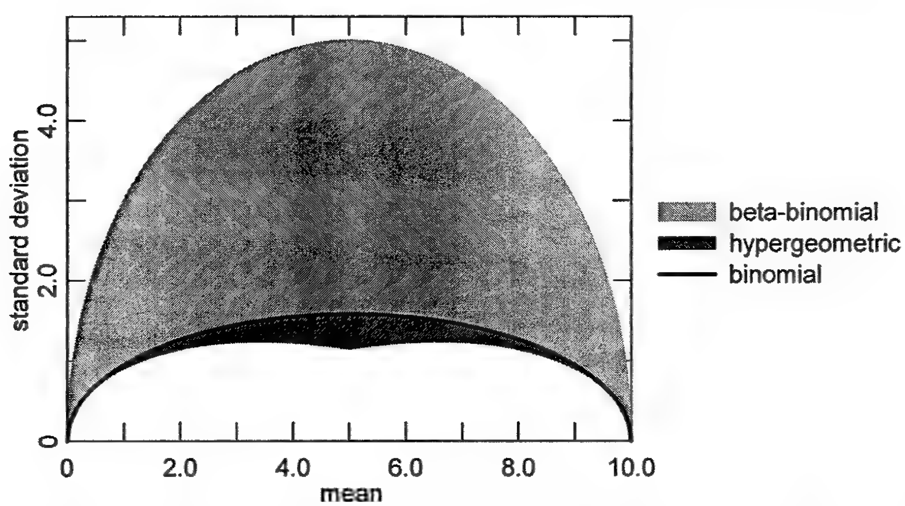

Figure 3.4: Relationship between standard deviation and mean for the three discrete distributions with n *= 10. The heavy line shows the dependence of standard deviation on mean for the binomial distribution; the shaded areas show the accessible values for the other distributions. The accessible region for the hypergeometric distribution is actually a lattice of isolated points, since the parameters of that distribution must be integers; the density of the lattice increases toward the heavy line.*

be integers. Since $N \to \infty$ corresponds to approaching the heavy line in Figure 3.4 from below, the lattice of points increases in density toward the heavy line.

All three distributions have positive skewness if $\mu < n/2$, zero skewness when $\mu = n/2$ and negative otherwise, as expected. The behaviour of the skewness and kurtosis of the binomial distribution is shown in Figure 3.2. Corresponding plots for the other two distributions are more complicated because their skewness and kurtosis do not depend on the coefficient of variation σ/μ alone. That is, η_3 and η_4 for the hypergeometric and β-binomial distributions are functions of two variables with domains of definition shown as the shaded areas in Figure 3.4. The β-binomial distribution is considered first, since it has a well defined parameter inversion (Eqn (3.28)).

Figure 3.5 shows contour plots of skewness and kurtosis for the β-binomial distribution. The broken lines in this Figure outline the domain of definition, which is the lighter-shaded area in Figure 3.4. Only the left half of the domain is shown; skewness is antisymmetric about $\mu = n/2$ and kurtosis is symmetric. As Figure 3.5 indicates, both skewness and kurtosis increase without limit as $\mu_{\beta\text{-b}} \to 0$. In terms of distribution parameters, this corresponds to $a \to 0$. The skewness is zero when $\mu_{\beta\text{-b}} = n/2$, that is, when $b = a$. The smallest kurtosis value also lies on this line and at the top of the domain of definition, which corresponds to the parameters $a \to 0$, $b \to 0$. The value of this minimum is 1.0, the smallest encountered in this work for any distribution, whether continuous or discrete, finite or infinite. In fact, as one traces the line $\mu_{\beta\text{-b}} = n/2$ from bottom ($a = \infty$) to top ($a = 0$), the kurtosis falls from the binomial-distribution value of $3 - n/2$ (= 2.8 for $n = 10$) to unity. That is, low kurtosis is found along the entire line given by $\mu_{\beta\text{-b}} = n/2$ and there is a region at the top of the domain of definition where the kurtosis is smaller than the value of 1.2 applying to the uniform distribution.

Figure 3.6 shows corresponding results for the hypergeometric distribution. This Figure was prepared using Equations (3.26) and taking $N = N$, $M = M$, that is, ignoring the discretisation of the domain. As with the β-binomial distribution, the skewness of the hypergeometric distribution is zero on the line $\mu_h = n/2$ and rises without limit as $\mu_h \to 0$. The inset indicates this rather schematically. (In terms of parameters, $\mu_h \to 0$ corresponds to $N \to \infty$.) The kurtosis also rises without limit as $\mu_h \to 0$. On the line $\mu_h =$

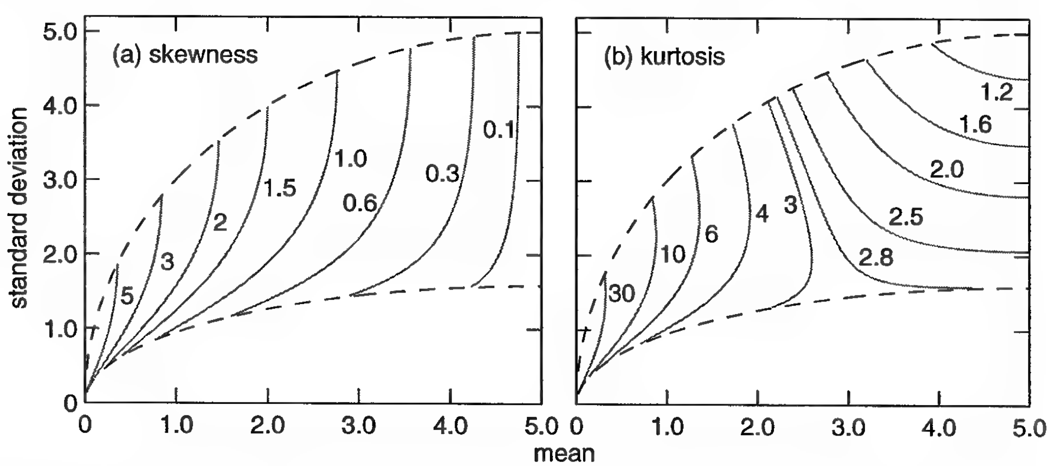

Figure 3.5: Contours of constant (a) skewness and (b) kurtosis for the β-binomial distribution with n = 10. *The broken lines delineate the domain of definition, which is the same as the lighter shaded area in Figure 3.4. Only the left half of the domain is shown; the skewness is antisymmetric about* $\mu_{\beta\text{-}b} = 5$, *the kurtosis symmetric.*

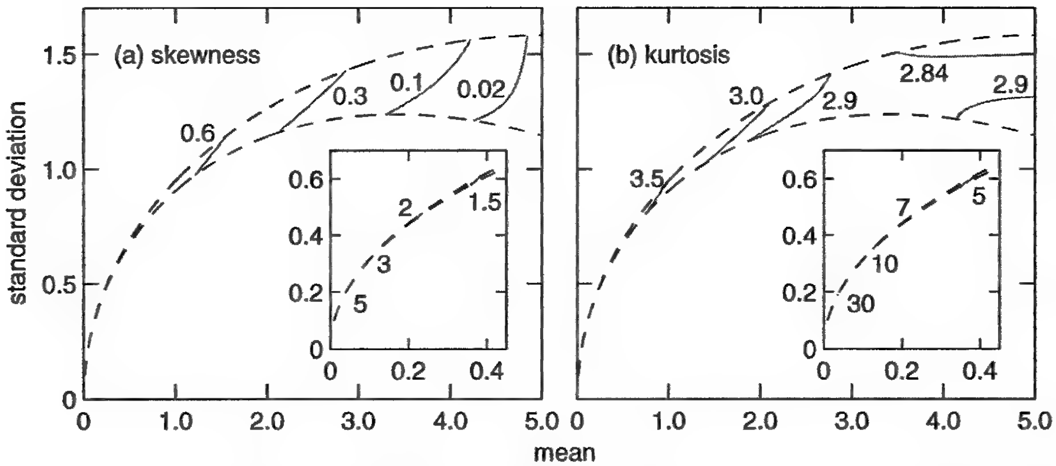

Figure 3.6: As in Figure 3.5, but for the hypergeometric distribution. In this case, the domain of definition is the darker shaded area in Figure 3.4. The inset in each panel shows detail near the origin. Because the upper and lower domain boundaries are so close to each other in this region, contours are not discernable in the insets; they lie near to the numbers.

$n/2$ (corresponding to $N = 2M$), the kurtosis is a minimum at the top of the domain, equalling the value for the binomial distribution (i.e. $3 - n/2$ or 2.80 for $n = 10$). The kurtosis rises slowly as one moves down along the line $\mu_h = n/2$.

3.3.6 Less Well Known Distributions

This Section briefly describes several other discrete distributions with a finite domain $[0,n]$. The study of some of these has a long history, but they nevertheless do not appear in most compilations of distributions. The purpose of this Section is to promote awareness of the distributions. The expression of their parameters in terms of their moments is carried through only where it is simple.

(a) Distributions with Zero Parameters

As explained in §3.3.1, frequently the range n of the distribution domain is set by external considerations; it is not available for variation. Hence, distributions with n as the only variable other than the argument are considered in this Report to have no parameters. Lest it be thought that the discrete uniform distribution is the only one such, three others are described here.

The 'classical matching distribution' is of some antiquity. It is the answer to the following question: n entities are numbered consecutively $1,\ldots,n$ and are then rearranged at random. What is the probability that t of them will have a position in the random sequence that matches their number? The result is [25(p.87),26(p.409)][(m)]

$$p_{\mathrm{cm}}(t) = \frac{1}{t!}\sum_{j=0}^{n-t}\frac{(-1)^j}{j!} \qquad (0 \le t \le n). \tag{3.32}$$

The moments of this distribution are: $\mu = 1$, $\sigma = 1$, $\eta_3 = 1$, $\eta_4 = 4$. This distribution is the finite-domain equivalent of the Poisson distribution, in the sense that the first n factorial moments of the two distributions are equal [26(p.410)].

(m) The equation in Ref. 26 (Eqn 10.19) contains a misprint in the lower limit of the sum.

Figure 3.7 shows the classical matching distribution with $n = 10$ (left-most bar in each group). The values of $p_{\mathrm{cm}}(0)$ and $p_{\mathrm{cm}}(1)$ are almost equal, after which p_{cm} values fall rapidly with increasing t, as the inset indicates. The last values are $p_{\mathrm{cm}}(n-1) = 0$ and $p_{\mathrm{cm}}(n) = 1/n!$.

Naor's distribution [26(p.447)] belongs to a type known as 'urn models' [24]. It gives the probability that t attempts are required to draw a red ball from a urn that initially contains one red ball and $n - 1$ white balls, under the condition that, every time a white ball is drawn, it is replaced with a red ball. Clearly, this sort of problem admits almost endless variation. As described above, the domain of Naor's distribution is $1 \le t \le n$. A straightforward change of variable gives

$$p_{\mathrm{N}}(t) = \frac{n!\,(t+1)}{(n-t)!\,(n+1)^{t+1}} \qquad (0 \le t \le n) \tag{3.33}$$

Despite its simple form, there are apparently no known closed-form expressions for the moments of this distribution. Figure 3.7 shows the distribution with $n = 10$ (right-most bar in each group). The distribution rises to a maximum at $n = 2$ and then falls, slowly at first. The last value is $p_{\mathrm{N}}(n) = n!/(n+1)^n$.

A more bizarre example of a parameterless distribution is 'Haight's harmonic distribution', which is [26(p.470)]

$$p_{\mathrm{Hh}}(t) = \frac{1}{2Z}\left(\left\lfloor \frac{2Z}{2t+1} \right\rfloor - \left\lfloor \frac{2Z}{2t+3} \right\rfloor\right) \qquad (0 \le t \le \lfloor Z - 1/2 \rfloor), \tag{3.34}$$

where $2Z$ is a positive integer and $\lfloor x \rfloor$ is the largest integer less than or equal to x. There does not seem to be any particular application behind the definition of this distribution. It has the property that, for large Z values, $p_{\mathrm{Hh}}(t)$ is zero for considerable ranges of t, with isolated nonzero values. Figure 3.7 shows something of this behaviour (middle bars of each group), although $n = 10$ is not large enough for it to be fully developed. Expressions for the mean and standard deviation of this distribution are available [26 (pp.470–1)].

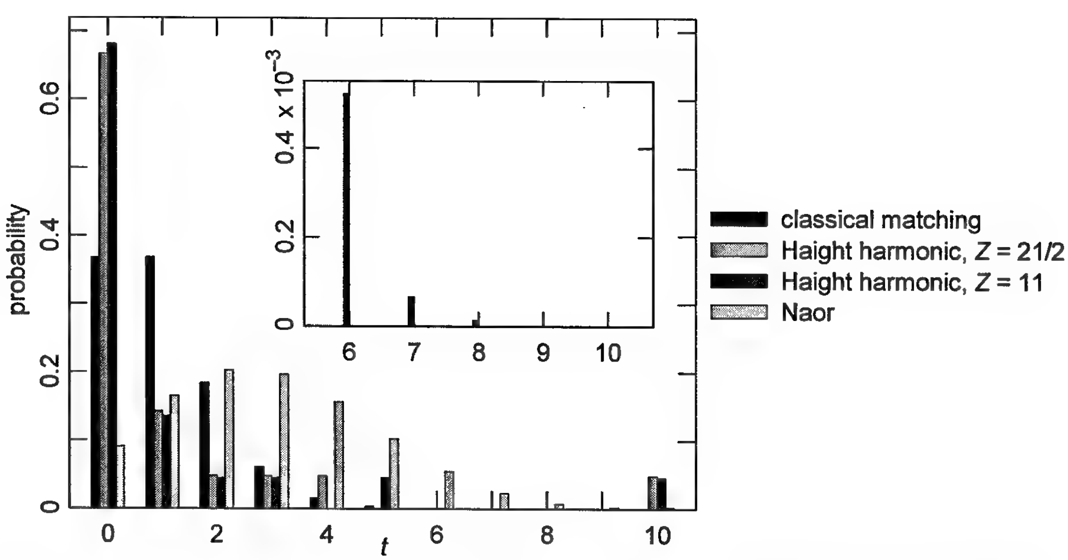

Figure 3.7: Four parameterless distributions with n = 10*. The inset shows detail of the classical matching distribution for 6 ≤*t*≤10. See text for further comments.*

It is not quite right to describe Haight's harmonic distribution as parameterless: if one wishes an upper limit of n, then $Z = n + ½$ and $Z = n + 1$ are both possibilities and give different distributions, as Figure 3.7 shows. (In this case—$n = 10$—the most pronounced difference is that $p_{\mathrm{Hh}}(4) = 0$ for $Z = 11$, but $p_{\mathrm{Hh}}(5) = 0$ for $Z = 21/2$.) In this sense, a parameter variation, albeit highly restricted, can be performed.

(b) One-Parameter Distributions

The 'riff-shuffle distribution' is related to a problem in combining two packs of cards. Its definition is [26(p.234)][(n)]

$$p_{\text{r-s}}(t) = \binom{n+t}{t}\left[a^{n+1}(1-a)^t + a^t(1-a)^{n+1}\right] \qquad (0 \le t \le n), \tag{3.35}$$

where $0 \le a \le 1$. This distribution is related to the negative binomial distribution. It is symmetric about $a = 0.5$ and tends to $\delta_{t,0}$ as $a \to 0$ or as $a \to 1$.[(o)] Expressions for the moments are available in terms of incomplete β functions [28]. Examples of the distribution for $n = 10$ are shown in Figure 3.8.

Dandekar introduced several distributions related to the binomial and Poisson distributions, in which the probability of a success in a trial is set to zero for a specified number of trials following a successful trial. His first modified Poisson distribution has one parameter only [25(p.25),26(p.435)]:

$$p_{\mathrm{D1P}}(t) = \mathrm{e}^{-\lambda(1-t/\nu)} \sum_{j=0}^{t} \frac{\lambda^j (1-t/\nu)^j}{j!} \qquad (0 \le t \le \lfloor \nu \rfloor), \tag{3.36}$$

where λ and ν are both greater than zero and $\lfloor x \rfloor$ is the largest integer less than or equal to x. To have any members in the distribution, $\nu \ge 1$. As with Haight's harmonic distribution, some variation of ν is available once the maximum range n has been chosen:

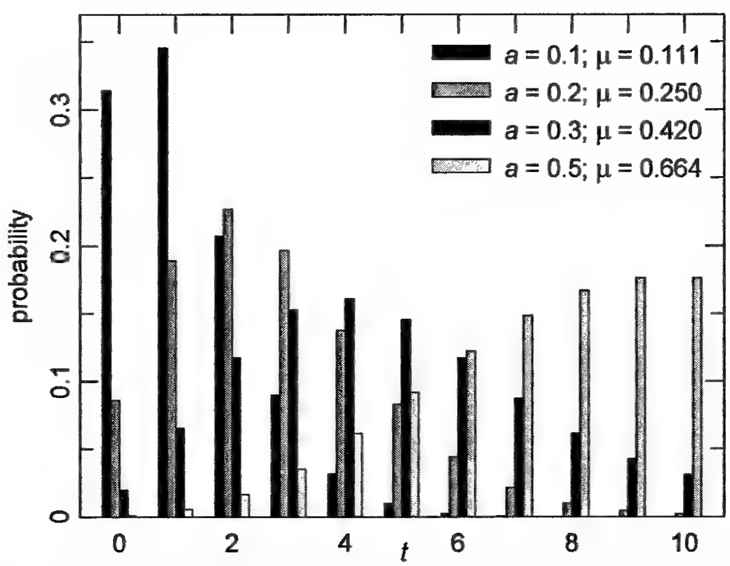

Figure 3.8: Examples of the riff-shuffle distribution for n *= 10. Calculated values of the mean are shown in the legend. The distribution tends to* $\delta_{t,0}$ *as* a $\to$ *0 or* a $\to$ *1.*

(n) The distribution is incorrectly stated by Johnson *et al.* [26(Eqn 5.92)], see Ref. 28.

(o) $\delta_{m,n}$ is the 'Kronecker delta': $\delta_{m,n} = 1$ if $m = n$ and zero otherwise.

$n \le \nu < n+1$, although in this case, the effect of varying ν over this range is relatively minor. As $\lambda \to 0$, the distribution degenerates to $\delta_{t,0}$. It seems that no simple expressions for the moments are known.

Note that Equation (3.36) gives the *cumulative* probability distribution. The usual probability distribution can be calculated by taking successive differences. Figure 3.9 shows some cases for $\nu = 10.0$. On the basis of this cursory numerical inspection, it appears that the distribution approaches $\delta_{t,n}$ as $\lambda \to \infty$.

The Bose-Einstein distribution comes from quantum statistical mechanics. Of the three state-occupancy distributions of statistical mechanics, it is the only one relevant here; the Maxwell-Boltzmann distribution is equivalent to the binomial distribution and the Fermi-Dirac distribution is defined for $0 \le t \le 1$ only. All three distributions are examples of urn models. Using notation consistent with that adopted in this Report, the Bose-Einstein distribution of state occupancies can be written [24(p.112),26(p.421)]

$$p_{\text{B-E}}(t) = \frac{(N-1)n!\,(n+N-t-2)!}{(n+N-1)!\,(n-t)!} \qquad (0 \le t \le n), \tag{3.37}$$

where $N \ge 2$ and is integral. This is one case where expressions for the mean and variance are simple. That for the mean, $\mu = n/N$, gives an expression for N, which leads to the variance as

$$\sigma^2 = \frac{\mu(n-\mu)(\mu+1)}{n+\mu}. \tag{3.38}$$

In statistical mechanics, n is the number of particles and N the number of states, both of which are always very large in physics applications, but the distribution is perectly well defined for small values also. Figure 3.10 shows four examples with $n = 10$.

The Bose-Einstein distribution becomes the discrete uniform distribution when $N = 2$; as $N \to \infty$, it degenerates to $\delta_{t,0}$.

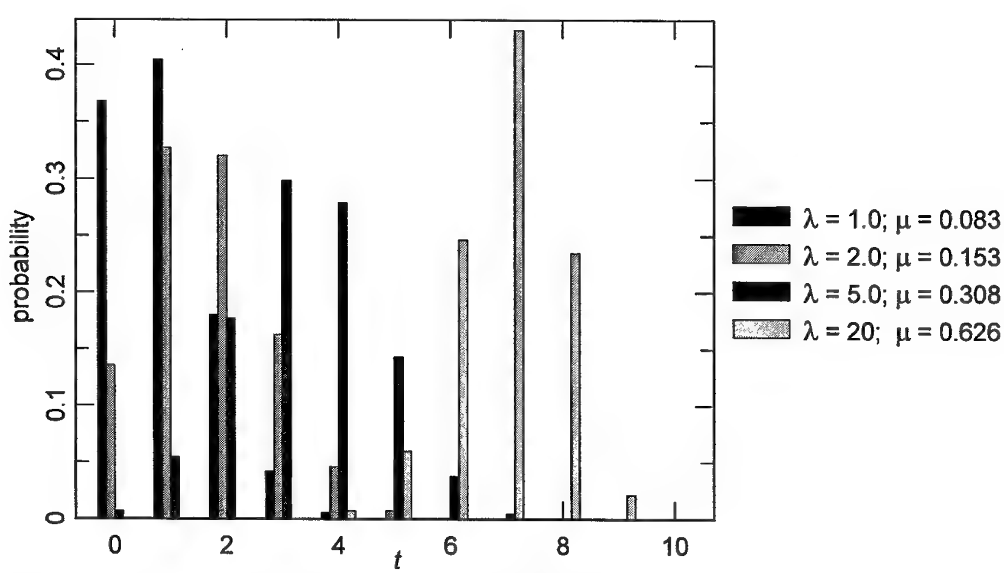

Figure 3.9: Examples of Dandekar's first modified Poisson distribution with $\nu = 10$. Calculated values of the mean are shown in the legend. The distribution tends to $\delta_{t,0}$ as $\lambda \to 0$ and $\delta_{t,10}$ as $\lambda \to \infty$.

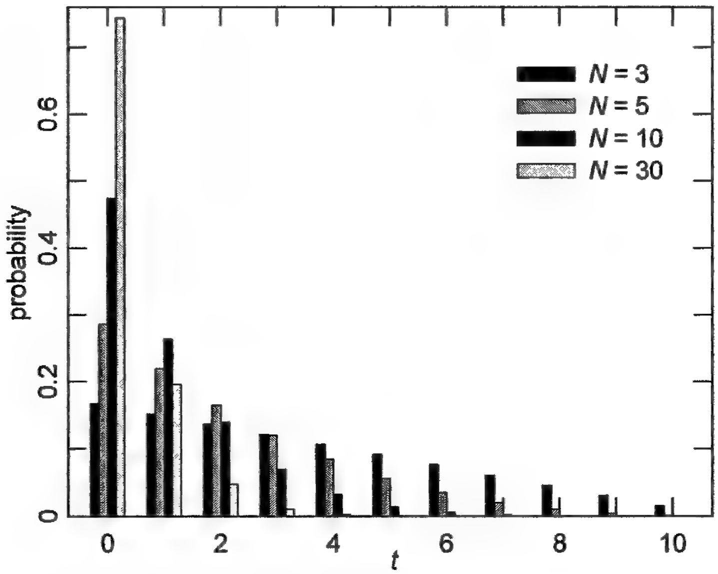

Figure 3.10: Examples of the Bose-Einstein distribution for n = 10. *The distribution equals the discrete uniform distribution when* N = 2 *and tends to* $\delta_{t,0}$ *as* N $\to \infty$.

(c) Distributions with Two or More Parameters

The Laplace-Haag matching distribution is a two-parameter generalisation of the classical matching distribution [26(p.410)]:

$$p_{\text{L-H}}(t) = \frac{n!}{N!\,t!} \sum_{j=0}^{n-t} \frac{(-1)^j (N-t-j)!}{j!\,(n-t-j)!} a^{t+j} \qquad (0 \le t \le n), \tag{3.39}$$

where $a > 0$ and $N \ge \max(n, na)$. In the original Laplace version of the distribution, a is integral and $N = na$. However, the distribution is well defined for any positive a, including $a < 1$.

A general expression for the factorial moments of the Laplace-Haag distribution is known [26(p.410)]. This leads to expressions for the mean and variance that can be readily inverted:

$$a = \frac{\mu\left(\mu^2 + \sigma^2 - \mu\right)}{n\sigma^2 + \mu^2 - n\mu}, \qquad N = na/\mu. \tag{3.40}$$

The condition $N \ge na$ on parameter N means that $\mu \le 1$, regardless of the value of n.

The 'binomial distribution of order k' is an example of a class of distributions that has been much studied in the last two decades. It is [26(p.431)]

$$p_{\text{b}k}(t) = p^n \sum_{j=0}^{k-1} \sum_{x_1} \cdots \sum_{x_k} \binom{x_1 + \dots + x_k + t}{x_1, \dots, x_k, t} \left(\frac{1}{p} - 1\right)^{x_1 + \dots + x_k} \qquad \left(0 \le t \le \left\lfloor \frac{n}{k} \right\rfloor\right), \tag{3.41}$$

where $0 \le p \le 1$,

$$\binom{x_1 + \dots + x_k + t}{x_1, \dots, x_k, t} = \frac{\left(t + \sum_{l=1}^{k} x_l\right)!}{t! \prod_{l=1}^{k} x_l!} \tag{3.42}$$

is the multinomial symbol and the inner summations in Equation (3.41) run over all

non-negative subscript values that satisfy

$$\sum_{l=1}^{k} l x_l = n - j - kt . \tag{3.43}$$

This is effectively a 2-parameter distribution. Despite the complication of its definition, relatively simple expressions for its mean and variance are known [26(p.431)], but they cannot be readily solved in favour of p and k.

The 'specified occupancy distribution' [26(p.416)] is a two-parameter generalisation of the classical occupancy distribution:

$$p_{\mathrm{so}}(t) = \frac{n!}{t!} \sum_{j=t}^{n} \frac{(-1)^{j-t}}{(n-j)!\,(j-t)!} \left(1 - \frac{j}{c}\right)^{b} \qquad (0 \le t \le n), \tag{3.44}$$

where $b > 0$ and $c \ge n$ are both integers. Once again, expressions for the moments are available but they cannot be inverted in closed form.

As a final example, the 'multinomial distribution', a generalisation of the binomial distribution, has as many parameters as one desires. Its definition is [25(p.104),26(p. 460)]

$$p_{\mathrm{m}}(t) = \sum_{r_0} \cdots \sum_{r_s} \binom{n}{r_0, r_1, \ldots, r_s} \prod_{j=0}^{s} p_j^{\,r_j} \qquad (0 \le t \le ns), \tag{3.45}$$

where the p_j, each separately satisfying $0 \le p_j \le 1$, also satisfy the third condition below and the summations run over all non-negative integers that satisfy the first and second following conditions:

$$\sum_{j=0}^{s} r_j = n , \qquad \sum_{j=0}^{s} j r_j = t , \qquad \sum_{j=0}^{s} p_j = 1 . \tag{3.46}$$

As a consequence of the last condition, only s of the $s + 1$ values of p_j are independent, but s itself can be considered a parameter, giving $s + 1$ parameters in all. The case $s = 1$ is the binomial distribution. Quite simple expressions for the moments are known, but the task of inverting $s + 1$ equations is substantial.

4. Conclusion

This report collects, summarises and extends the knowledge of the properties of skewed probability distributions, with the chief goal of providing a means of comparing among them. The motivation for this was drawn from various studies in operational analysis in which random variables are employed without any firm information on their probability distributions. In this situation, it is of interest to look for sensitivity to the distribution chosen. This can best be done by running the model with a variety of distributions, which then raises the question of how to compare them. The work began with the premise that a viable method for such comparison is to match distribution moments, but this requires access to the requisite equations. Surprisingly, these are almost completely absent from compilations of information on probability distributions. Hence, the main task addressed in this report—and addressed successfully—is the derivation and compilation of the required equations for common skewed distributions. A total of 18 distributions are treated in detail, both finite and semi-infinite,

discrete and continuous. These include three non-skewed distributions, for comparison. An additional 11 discrete distributions are briefly mentioned.

In the process, it is noted that most distributions have a limited accessible range of the coefficient of variation (the ratio σ/μ of standard deviation to mean). A significant part of this work is the determination of these limits. The results are presented above for each distribution; they are also collected in Table 2.5, Table 3.1 and Figure 3.4. The limits typically place limits on accessible values of the higher moments, which are the main indicators of difference between distributions with equal mean and standard deviation.

The distributions treated in this report comprise all of the most commonly used skewed distributions. The collected formulae and graphs should be of use to anyone using skewed random variates who needs a common and consistent basis for comparing one distribution with another.

References

[1] M.P. Fewell & M.G. Hazen (2004) 'Cognitive issues in modelling network-centric command and control', submitted for publication as a Research Report of the Defence Science and Technology Organisation.

[2] M.P. Fewell & I. Grivell (2004) 'Effect of service-time probability distribution on the behavior of queues dominated by reneging' submitted for publication in *Mil. Op. Res.*

[3] R. Klingbeil (2001) 'An analysis of command decision time delay' in *Proc. 1st meeting of MAR Group AG-1*, TTCP document DOC-MAR-1-2001, vol. 2.

[4] R. Klingbeil & K. Sullivan (2003) *An analysis of decision-making time delay*, Technical memo NUWC-NPT TM 03-028 of the Naval Undersea Warfare Center Division Newport, RI.

[5] I. Olkin, L.J. Gleser & C. Derman (1980) *Probability models and applications*, New York, MacMillan.

[6] S. Kokosa & C. Nevison (1988) *Statistical tables and formulae*, New York, Springer-Verlag.

[7] A.M. Law & W.D. Kelton (2000) *Simulation modeling and analysis*, 3rd edn, Boston, McGraw-Hill.

[8] M. Abramowitz & I.A. Stegun (1968) *Handbook of mathematical functions* (5th Dover printing), New York, Dover.

[9] M.C.K. Tweedie (1957) 'Statistical properties of the inverse Gaussian distribution', *Ann. Math. Stat.* **28**, 362–77 and 696–705.

[10] N.L. Johnson, S. Kotz & N. Balakrishnan (1995) *Continuous univariate distributions*, 2nd edn (2 vols.), New York, John Wiley

[11] R.S. Chhikara & J.L. Folks (1989) *The inverse Gaussian distribution*, New York, Marcel Dekker.

[12] M. Evans, N. Hastings & B. Peacock (1993) *Statistical distributions*, 2nd edn, New York, John Wiley.

[13] K. Bury (1998) *Statistical distributions in engineering*, Cambridge, CUP.

[14] B.R. Haverkort (1998) *Performance of computer communication systems: a model-based approach*, Chichester, John Wiley.

[15] H.B. Dwight (1971) *Tables of integrals and other mathematical data*, 4th edn., New York, Macmillan.

[16] W.P. Elderton and N.L. Johnson (1969) *Systems of frequency curves*, Cambridge, CUP.

[17] A.J. Gross (1983) 'Hazard plotting' and J.K. Patel (1983) 'Hazard rate and other classifications' in S. Kotz, N.L. Johnson & C.B. Read (eds) *Encyclopædia of Statistical Sciences*, vol. 3, New York, Wiley Interscience.

[18] D.H. Olwell (1997) 'Modeling loss exchange ratios as inverse Gaussian variates: implications', *Mil. Op. Res.* 3(1) 51–67.

[19] R.E. Barlow & F. Proschan (1981) *Statistical theory of reliability and life testing – probability models*, 2nd corrected printing, Silver Spring MD, To Begin With.

[20] R.N. Bracewell (1978) *The Fourier transform and its applications* (2nd edn), Tokyo, McGraw-Hill.

[21] A.A.K. Ayyangar (1941) 'The triangular distribution', *Math. Student* **9**, 85–7.

[22] F.A. Haight (1978) 'Special discrete distributions' in W.H. Kruskal & J.M. Tanur (eds) *International Encyclopædia of Statistics* (2 vols), New York, The Free Press.

[23] M. Sibuya (1983) 'Generalised hypergeometric distributions' in S. Kotz, N.L. Johnson & C.B. Read (eds) *Encyclopædia of Statistical Sciences*, vol. 3, New York, Wiley Interscience.

[24] N.L. Johnson & S. Kotz (1977) *Urn models and their applications*, New York, John Wiley.

[25] G.P. Patil, M.T. Boswell, S.W. Joshi & M.V. Ratnaparkhi (1984) *Dictionary and classified bibliography of statistical distributions in scientific work, vol. 1: discrete models*, Fairland MD, International Cooperative Publishing House

[26] N.L. Johnson, S. Kotz & A.W. Kemp (1992) *Univariate discrete distributions*, 2nd edn, New York, Wiley.

[27] W.C. Guenther (1975) 'The inverse hypergeometric—a useful model', *Stat. Neerl.* **29**, 129–44.

[28] V.R Uppuluri & W.J. Blot (1970) 'A probability distribution arising in a riff-shuffle', in G.P. Patil (ed.) *Random counts in scientific work 1: Random counts in models and structures*, University Park PA, Pennsylvania State University Press, pp. 23–46.

DISTRIBUTION LIST

Comparative Descriptive Statistics of Skewed Probability Distributions

M.P. Fewell

AUSTRALIA

DEFENCE ORGANISATION

Task Sponsor	Director General Navy Strategic Policy and Futures	1
	Director Navy Strategy and Futures	1
	Deputy Director Future Maritime Warfare	1

S&T Program

Chief Defence Scientist	shared copy
FAS Science Policy	shared copy
AS Science Corporate Management	shared copy
Director General Science Policy Development	shared copy
Counsellor Defence Science, London	Doc Data sheet
Counsellor Defence Science, Washington	Doc Data sheet
Scientific Adviser to MRDC, Thailand	Doc Data sheet
Scientific Adviser Joint	1
Navy Scientific Adviser	1
Scientific Adviser - Army	Doc Data sheet
Air Force Scientific Adviser	Doc Data sheet
Scientific Advisor to the DMO M&A	Doc Data sheet
Scientific Advisor to the DMO ELL	Doc Data sheet

Systems Sciences Laboratory

Director System Sciences Laboratory	Doc Data sheet
Chief of Maritime Operations Division	1
Research Leader Maritime Combat Systems	1
Research Leader Maritime Operations Research	1
Head Maritime Tactical Experimentation, MOD	1
Head Maritime Operations and Tactics Analysis, MOD	1
Head Naval Capability Studies, MOD	1
Head Undersea Warfare Analysis, MOD	1
Dr A. Knight, MOD	1
Dr J. Smelt, MOD	1
Dr M.P. Fewell, MOD	15

Information Sciences Laboratory

Dr. I. Grivell, IND	1

OR Hub Executive

Dr N. Curtis, RLOAE	1
Mr G. Murray, AOD	1
Mr J. Smith, RLAOA	1
Dr P. Whitbread, RLTCA	1
Dr D. Wood, RLSCS	1

DSTO Library and Archives	
Library Edinburgh	2
Australian Archives	1
Capability Systems Staff	
Director General Maritime Development	Doc Data sheet
Director General Information Capability Development	Doc Data sheet
Office of the Chief Information Officer	
Deputy Chief Information Officer	Doc Data sheet
Director General Information Policy and Plans	Doc Data sheet
AS Information Strategy and Futures	Doc Data sheet
AS Information Architecture and Management	Doc Data sheet
Director General Australian Defence Simulation Office	Doc Data sheet
Strategy Group	
Director General Military Strategy	Doc Data sheet
Director General Preparedness	Doc Data sheet
HQAST	
SO (Science) (ASJIC)	Doc Data sheet
Navy	
SO (Science), COMAUSNAVSURFGRP, NSW	Doc Data sheet & dist list
Director General Navy Capability, Performance and Plans, Navy HQ	Doc Data sheet
Air Force	
SO (Science), Headquarters Air Combat Group, RAAF Base, Williamtown	Doc Data sheet & Exec Summ
Army	
ABCA National Standardisation Officer, Land Warfare Development Sector, Puckapunyal	emailed Doc Data sheet
SO (Science), Deployable Joint Force Headquarters (DJFHQ) (L), Enoggera	Doc Data sheet
SO (Science) - Land HQ, Victoria Barracks NSW	Doc Data sheet & Exec Summ
Intelligence Program	
Director General Scientific and Technical Analysis, Defence Intelligence Organisation	1
Manager, Information Centre, Defence Intelligence Organisation	1 printed & 1 pdf
AS Corporate, Defence Imagery and Geospatial Organisation	Doc Data sheet
Defence Materiel Organisation	
Head Aerospace Systems Division	Doc Data sheet
Chief Joint Logistics Command	Doc Data sheet
Defence Libraries	
Library Manager, Defence Library Service - Canberra	Doc Data sheet
Library Manager, Defence Library Service - Sydney West	Doc Data Sheet

UNIVERSITIES AND COLLEGES	
Library, Australian Defence Force Academy	1
Head Aerospace and Mechanical Engineering, Australian Defence Force Academy	1
Serials Section (M list), Deakin University Library	1
Hargrave Library, Monash University	Doc Data sheet
Librarian, Flinders University of SA	1
Dr. G. Yeo, School of Mathematics and Statistics, University of Western Australia	1
OTHER ORGANISATIONS	
National Library of Australia	1
NASA (Canberra)	1

OUTSIDE AUSTRALIA

INTERNATIONAL DEFENCE INFORMATION CENTRES	
US Defense Technical Information Center	2
UK Defence Research Information Centre	2
Canada Defence Scientific Information Service	1
NZ Defence Information Centre	1
ABSTRACTING AND INFORMATION ORGANISATIONS	
Library, Chemical Abstracts Reference Service	1
Engineering Societies Library, US	1
Materials Information, Cambridge Scientific Abstracts, US	1
Documents Librarian, The Center for Research Libraries, US	1
THE TECHNICAL COOPERATION PROGRAM	
Maritime Systems Group	
Chair, Action Group 1	1
Canadian National Leader, Action Group 1	1
New Zealand National Leader, Action Group 1	1
United Kingdom National Leader, Action Group 1	1
United States National Leader, Action Group 1	1
Dr R. Klingbeil, Member, Action Group 1	1
Mr Mark G. Hazen, Member, Technical Panel 1	1
SPARES	5

Total number of copies: 70

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION
DOCUMENT CONTROL DATA

1. PRIVACY MARKING/CAVEAT (OF DOCUMENT)
None

2. TITLE Comparative Descriptive Statistics of Skewed Probability Distributions	3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION) Document (U) Title (U) Abstract (U)
4. AUTHOR(S) M.P. Fewell	5. CORPORATE AUTHOR DSTO System Sciences Laboratory, P.O. Box 1500, Edinburgh, SA 5111, Australia

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-TR-1596	AR-013-136	Technical Report	July 2004

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
E 9505-28-9	NAV 01/389	DGNSPF	37	28

13. DOWNGRADING/DELIMITING INSTRUCTIONS	14. RELEASE AUTHORITY
N.A.	Chief, Maritime Operations Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111

16. DELIBERATE ANNOUNCEMENT

No limitations

17. CITATION IN OTHER DOCUMENTS Yes

18. DEFTEST DESCRIPTORS

Probability distribution functions, Skewed density functions, Discrete distribution functions, Continuous distribution functions, Random variables

19. ABSTRACT

This report is a brief handbook on the comparative descriptive statistics of a wide variety of skewed probability distributions, both continuous and discrete. The aim is to facilitate the comparison of different distributions, for use where random variables are employed without any firm information on their distribution. In this situation, it is of interest to look for sensitivity to the distribution chosen. This can best be done by running the model with a variety of distributions, which then raises the question of how to compare distributions. This work advocates the use of moments and presents the requisite equations. As obvious as this approach may appear, many of the equations do not seem to have been published previously and some of the results are apparently wholly new. A total of 18 distributions are treated in detail, including all of the most commonly used skewed probability distributions.